Journal of Urban Economics 134 (2022) 103514

Contents lists available at ScienceDirect

Journal of Urban Economics

journal homepage: www.elsevier.com/locate/jue

The price paid: Heuristic thinking and biased reference points in the housing market

Charlotte C. Meng

Queen Mary University of London, United Kingdom

A R T I C L E I N F O

JEL codes:
R31
D91
D12

Keywords:
Left-digit Bias
Reference dependence
Behavioural economics
Housing markets

A B S T R A C T

Does the power of reference points mean that minute differences in a purchase price then reverberate in future sales prices? In this research, I show that if previous sales prices are round numbers, defined as multiples of £1,000 (e.g. £231,000), subsequent sales prices entail a considerable premium relative to similar properties that were previously priced at charm numbers that are marginally below those round numbers (e.g. £230,999 or £230,950). Using a sample of repeat sales from the Greater London region from 1995 to 2017, I estimate the premium to be approximately 4 percent after controlling for property characteristics and a large set of fixed effects. Increasing public accessibility of information attenuates the effect. Tax considerations, financial constraints, and pricing errors cannot explain the result. I propose a framework of reference dependence and left-digit bias to explain the result, highlighting the presence of behavioural biases in household decisions, even when very high stakes are involved.

1. Introduction

A home is typically the most valuable household asset, which means that sellers are motivated to maximise the price at time of sale. Home sellers in repeat sales are prone to reference dependence bias. This means they frequently refer to the price they paid to acquire the house as a reference point and are reluctant to sell at prices below this reference point because of the feeling of loss it may cause (Anenberg, 2011; Engelhardt, 2003; Genesove and Mayer, 2001; Kahneman, 1979; Tversky and Kahneman, 1992). In this research, I revisit the *price paid* as the reference point, drawing on the growing evidence about people's inattention to transaction details from recent psychological and economic studies. Evidence shows that inattention affects people's perceptions of previous sales prices, distorts reference prices, and converts minute differences in prices paid in the past to considerable gaps in sales prices achieved in the future.

Attention is a limited resource and is used selectively — few people pay attention to every facet of something they are buying. This can leave people vulnerable to heuristics — or cognitive short cuts — when making purchase decisions. One form of inattention is left-digit bias, when people faced with a multi-digit number only focus on a few left-most digits rather than all digits to determine its value (Korvorst and Damian, 2008; Poltrock and Schwartz, 1984). It usually leads to round numbers, i.e. numbers with endings of multiple zeros, being perceived equal to or even higher than their true values. In contrast, charm numbers, i.e. numbers just below round numbers, are perceived as lower than their actual values. The bias is likely to have an especially large effect in the real estate market, where home prices are typically in the six-digit range, meaning that the amounts of money being overlooked are potentially large.

To facilitate the interpretation of the empirical findings, I first develop a framework of repeat sales in the housing market with reference dependence and left-digit-biased reference points. It predicts that properties with round-numbered previous prices are sold subsequently at a premium when compared with similar properties with charm-numbered prior prices. Suppose there are two similar properties *A* and *B*. *A* was previously sold at £231,000 and *B* at £230,950. Although £231,000 is only marginally (0.002 percent) higher than £230,950, it has a higher number in the thousands' column. An inattentive reader who only refers to the left three digits will perceive £231,000 to be a lot greater than £230,950. This is primarily the reason why the then-buyer of property B paid £230,950. In the subsequent sale, this discontinuous increase at round number thresholds in perceived prices passes on to the seller's reservation price, the buyer's willingness-to-pay, and ultimately the final transaction price. In addition, more rounded thresholds (e.g. £230,000) exhibit larger discontinuous increases than less rounded thresholds (e.g. £231,000). The then-buyer could have easily made the price £230,000 with an extra £50 to enjoy a better chance to obtain a higher price in the future. Missing the future benefit of this extra £50 is the *price paid* for inattention.

The model further states that time-on-market contains information that is useful for identifying the source of the bias, namely, whether the

E-mail address: c.meng@qmul.ac.uk

https://doi.org/10.1016/j.jue.2022.103514
Received 13 January 2021; Received in revised form 25 October 2022

discontinuity in transaction prices comes from distortions of the buyer's valuation, or the seller's valuation, or both. If left-digit bias only affects the seller, we would expect round-numbered properties to stay longer on the market, sellers waiting to obtain a premium. By contrast, if left-digit bias only distorts the buyer's evaluation, we would expect a shorter time-on-market for round-numbered properties. If it distorts both, then the effect on time-on-market is ambiguous. Section 7 utilises this prediction to test whose inattention is driving the results.

To test the predictions empirically, I compare the transaction prices of properties that are sold at a round number with those of properties that are sold at a number just below (i.e. a charm number). The method is similar to a regression discontinuity design. A large number in the left-most digits of the previous price is a treatment. A round price and its neighbouring charm prices are similar in values, but they differ in one or more left-most columns that have higher column values. Hence, properties with round-numbered previous prices are in the treatment group, and properties with charm-numbered previous prices neighbouring the exact round number are in the control group. In the neighbourhood of each round number, a control group and a treatment group form a pair. In the previous example, property A (£231,000) is in the treatment group and property B (£230,950) is in its paired control group. The fact that the two groups differ in their subsequent sales price would show that the previous purchase price plays a non-negligible role and is distorted by left-digit bias.

The analysis utilises the universe of house transactions in the Greater London region from 1995 to 2017. The empirical analysis compares, in multiples of £1000, properties sold at round numbers with those sold within £100 below that round number. The average subsequent sales prices have a 4 percent premium in round-numbered properties after controlling for property characteristics and an extensive set of fixed effects. The size of the effect grows with the roundness level. For 10,000-multiples, 50,000-multiples, and 100,000-multiples, the effect sizes are 4.3 percent, 5.5 percent, and 6.5 percent, respectively. The effect size falls over time, from 8.1 percent (before 2000) to only 4.1 percent (2008 onwards), which is consistent with the prediction of inattention theory that the salience of information attenuates the bias.

I evaluate the robustness of the finding and rule out several alternative explanations. First, stamp duty land tax (SDLT) in the UK was not smoothly defined by property values before December 2014. Thus, house prices around tax thresholds were distorted and bunched extensively under or at tax thresholds (Best and Kleven, 2018). In Section 6.1, a doughnut approach, whereby I remove all transactions around tax thresholds is utilised to minimise the estimation bias. My results suggest that tax considerations do not drive the price premiums.

Another possible explanation is that properties previously transacted at round prices have systematically tighter financial constraints, hold higher reservations prices, and eventually obtain higher transaction prices than those previously transacted at round prices (Anenberg, 2011; Genesove and Mayer, 1997; Stein, 1995). In Section 6.2, I estimate a loan-to-value (LTV) ratio at a postcode-level on a quarterly basis to gauge sellers' financial constraints and rule out this possibility.

Mispricing in previous prices – for example, price premiums related to the charm pricing strategy – is another potential explanation. In Section 6.3, I conduct a placebo test by investigating the transaction prior to the previous purchase. The identified effect cannot be explained by the potential overpricing component in charm prices alone (Allen and Dare, 2004; Chava and Yao, 2017; Repetto and Solís, 2020).

A natural question following the empirical finding is whose inattention is driving the results, the buyer's or seller's? To investigate this issue further, I bring in additional data about property listings including list prices, time on the market, and online page views. Similar discontinuity patterns are not only present in list prices, but more pronounced than those in the final sales price (4.3 percent versus 3.9 percent). This indicates that left-digit bias initially affected sellers' decision about list prices, confirming the reference dependence explanation. The two groups of properties do not differ in terms of time on the market or the number of online page views, controlling for list prices format, property characteristics, and fixed effects. According to the hypothesis in the theoretical model, the result implies that both buyers and sellers are affected by the bias.

My research contributes primarily to four strands of literature. First, it contributes to the recent literature about the role of numbers in decision making. While the magnitude of a number naturally attracts attention, its format also plays a role. Examples include the prevalence of round numbers as reference points, goals, or focal points for marathon runners (Allen et al., 2017), for home sellers and buyers (Pope et al., 2015), for consumers in a pay-what-you-want purchase (Lynn et al., 2013) and for scholastic assessment test (commonly known as SAT) takers in the US (Pope and Simonsohn, 2011). Another special format is a charm number that ends in one or more nines (e.g. 19.99, 299). Charm numbers are widely used in pricing and marketing strategies. According to a study by Holdershaw et al. (2017), 60 percent of advertised prices in newspapers end in the digit 9, 30 percent of prices ended in 5, 7 percent ended in 0, and only 3 percent end in the remaining seven digits. In recent years, empirical evidence has shown that charm pricing is related to better outcomes in retail (Shlain, 2018), negotiations (Hukkanen and Keloharju, 2019), and house transactions (Allen and Dare, 2004; Chava and Yao, 2017; Thomas et al., 2010). In this vein, this research adds empirical evidence of the numerical format's influences on house transaction outcomes.

Second, this research is closely related to a small body of work studying the reference point formation process. The location of the reference point is frequently assumed to be the *status quo*, or one's current assets. In their original paper about reference dependence, Kahneman and Tversky (1979) recognise that reference points are sometimes under the influence of various other factors, such as an expectation or aspiration level, and hence differ from the *status quo*. Existing research has made great endeavours to explore formation processes (Freeman, 2019; Kahneman et al., 1990; Koszegi, 2006; Maltz, 2019; Thaler, 1980). For repeat house sales, previous sales prices are considered a natural reference point (Anenberg, 2011; Engelhardt, 2003; Genesove and Mayer, 2001). This research contributes to this line of research by demonstrating that the reference point is not static and can be distorted by other heuristics such as left-digit bias.

Third, this research is also closely related to the growing literature on inattention. Despite the well-established research in psychology, inattention has only attracted the attention of economists in the recent decade. Recent literature has expanded the topic to various contexts. In finance, traders systematically underreact to Friday announcements (DellaVigna and Pollet, 2009). In marketing, consumers neglect the shrouded attributes of non-transparent tax (Chetty et al., 2009), shipping costs (Brownet al., 2010), and information that is intentionally undisclosed by sellers (Jin et al., 2015). The existing evidence is mostly from repeated decisions that involve small stakes. People may be inattentive in grocery shopping intentionally, because the event occurs frequently and the benefit from extra attention is usually low. However, for transactions that are high-stakes and infrequent, people are expected to be attentive to details. Evidence is lacking as to whether this is the case in real life.

Two such decisions that have been investigated are car purchasing and home purchasing decisions. Busse et al. (2013) and Lacetera et al. (2012) document that buyers in the car market are inattentive to mileage information, leading to discontinuous and considerable decreases in prices at round-numbered mileages. In the housing market, relevant empirical studies have explored left-digit bias and strategic listings at charm prices (Allen and Dare, 2004; Chava and Yao, 2017; Repetto and Solís, 2020). They show that properties that are listed at charm prices achieve higher sales prices than those listed at round prices. The present research differentiates from these other studies in two major aspects. First, it draws on the reference dependence feature of repeat sales so that the effects of unobserved housing attributes on house prices can be controlled. Second, this research focuses on the bias arising from both buyers' and sellers' evaluations, whereas these other studies focus on buyers' evaluations.

Finally, my research is related to the literature on house price determination. Traditional economic theory considers house prices to be rational and determined by a number of characteristics such as structure and location. Empirical studies in this direction have shown the effect of behavioural biases on house prices including loss aversion (Anenberg, 2011; Engelhardt, 2003; Genesove and Mayer, 2001) and anchoring (Arbel et al., 2014; Bucchianeri and Minson, 2013; Liu et al., 2015). This research contributes to this line of research by refining the reference dependence feature of repeat sales and documents that its effect is more than marginal.

In the late stage of revising this research, thanks to one of the anonymous referees, I became aware of a parallel work by Wiltermuth et al. (2021) using US data which shows empirically that previous sales prices matter for subsequent sales disproportionally around round numbered prices. The key difference between this present research and that of Wiltermuth et al. (2021) is that sales mechanisms in the two markets are different, which leads to different interpretations of the results. In the US, historical prices are not readily available during the time period examined, as highlighted by the authors in section 3.4. While sellers and their agents are aware of the historical prices, buyers do not know the prices. Therefore, inattention only affects seller-side valuation. In the UK, price information has become increasingly accessible to the public during the time period studied. This setting is therefore valuable in investigating both sellers' and buyers' valuations. In terms of empirical strategies, this research also differs from that of Wiltermuth et al. (2021) in several ways. First, this research is able to call upon a larger dataset, comprising of more than 200,000 repeat sales compared with 16,111 in Wiltermuth et al. (2021). A large dataset ensures that results are not driven by extreme prices and that findings are robust. Second, rather than a simple linear regression of the price, I utilise the repeat sales approach which is especially helpful in controlling for time-invariant unobservable characteristics. Finally, thanks to the large dataset, I have enough observations to estimate the inattention effect for multiple roundness levels, i.e. £1,000, £10,000, and £100,000 multiples, as detailed in section 5.3.

The remainder of the paper is structured as follows. Section 2 presents a simple model of inattention and biased reference points for repeat sales in the housing market, with two testable hypotheses. Section 3 describes the data followed by non-parametric and parametric analyses in section 4 and section 5, respectively. Section 6 provides some alternative explanations and robustness checks. Section 7 further investigates the potential mechanism behind the main results and brings in additional data to provide insights. Section 8 concludes the paper.

2. Theoretical framework

2.1. Biased reference points

In the second-hand housing market, previous prices serve as a salient reference point in both buyers' and sellers' decision-making process. This effect on sellers can be attributed to loss aversion (Genesove and Mayer, 2001). Sellers paid the price when they bought the property. So when current market prices are under previous prices, they tend to have feelings of loss, which motivates them to set the list price at a higher-than-market-price value. On the other hand, when current market prices are above previous prices – as is the case in most of our sample – sellers experience a feeling of gain. They are therefore willing to lower reservation prices and accept a price that they otherwise should not accept (i.e. without this reference dependence bias)[1].

The home buyer's evaluation is also affected by property price history, when it is accessible, in three possible ways. First, buyers consider previous prices as an initial signal of quality (Erdem et al., 2008; Fluet and Garella, 2002; Stiving, 2000). They are more likely to be attracted to view properties that were previously sold at higher prices, assuming that these properties are higher in quality or have positive qualities that are not reflected in the advertisements. More viewers mean more bidders in the auction and higher final transaction prices. Second, Tversky and Kahneman (1974) points out that people's decision-making process is usually affected by arbitrary values that they refer to as "anchors". Their ability to move away from those anchors is restricted. Buyers carry in mind the previous sold price when choosing bid prices for a property: the higher the previous price, the higher their bid prices. Third, buyers also anchor to list prices, as is shown the positive relationship between list prices and sales prices (Bokhari and Geltner, 2011; Bucchianeri and Minson, 2013). If sellers' reference-dependence behaviour inflates list prices, some of the premium will pass on to the final sales price through buyer's anchoring behaviour. In sum, all three effects predict that buyers' willingness-to-pay is increasing with reference prices.

Left-digit bias distorts decision maker's perception of the original purchase price, consequently the actual perceived reference prices exhibit discontinuous increases crossing each round-numbered threshold. This is because people read numbers from left to right and have the tendency to read some right digits with inattention, i.e. the perceived reference price $\hat{r} = r_{LDigits} + (1-\theta) r_{Remainder} = r - \theta \times r_{Remainder}$, where $r_{LDigits}$ represents the value of the left digits that people perceive with full attention. $r_{Remainder}$ represents the remaining digits that are perceived with limited attention; θ characterises the degree of inattention, such that $\theta = 0$ represents full attention and $\theta = 1$ represents complete ignorance. Assume that $\theta > 0$. Perceived values of round numbers, e.g. multiples of 1,000, equal their true values because $r_{Remainder} = 0$. By contrast, the numbers just below those round numbers, that is, charm numbers, have large $r_{Remainder}$ and consequently large perception bias, creating the discontinuity in perceived values crossing each round numbers.

2.2. The market equilibrium

I assume a simple linear utility function to demonstrate how reference prices affect equilibrium sales prices and sales time, as well as the outcomes caused by the left-digit bias. Suppose potential sellers get utility v_s if they do not sell their homes this period and get $p + \lambda(p-r)$ for selling their house for price p, where r is the reference price and λ captures the level of the bias. Then they would sell if $p + \lambda(p-r) \geq v_s$. The reservation price will be $p_s^* = \frac{1}{1+\lambda} v_s + \frac{\lambda}{1+\lambda} r$. It is increasing in r. Suppose a buyer's utility from buying the house at price p is $v_b - p - \gamma(p-r)$, and 0 otherwise. Then the most she will be willing to pay is $p_b^* = \frac{1}{1+\gamma} v_b + \frac{\gamma}{1+\gamma} r$.

Suppose a buyer comes along, with v_b and just bids p_b^*. Then the final price conditional on a sale is unambiguously increasing in r, with discontinuous effects for crossing round numbers (more so for more rounded numbers) as is shown in Section 2.1. This gives the first prediction: *properties with round previous prices have higher subsequent sales prices than similar properties with previous prices at neighbouring charm values; the more rounded the previous prices, the higher the price premium in subsequent sales, everything else being equal.*

If a sales price of £300,000 are worth much more to buyers (in terms of future value) than £299,999, then round prices should be far more common equilibrium outcomes. Buyers should just throw in the extra pound to get the future benefit. Why are there still charm sales prices? A possibility is that many buyers are not aware of left-digit bias and

[1] The effect does not necessarily lead to bunching at or around previous prices because of changes in market price levels, but instead adjustments in the direction of the previous prices: higher reference points induce a weaker feeling of gain or a stronger feeling of loss, which then leads sellers to adjust reservation prices upward.

don't anticipate this future sales effect. They like be able to say that they paid less than £300,000 for the house. The left-most digits rather than the actual value enters directly into their utility function.

In addition, assume that v_b is distributed according to a cumulative distribution function F. The probability of a sell not happening is equal to $\Pr(p_s^* > p_b^*)$. Plugging in the expressions for both reservation prices and simplifying, this probability is equal to $Pr(v_b < \frac{1+\gamma}{1+\lambda} v_s + \frac{\lambda-\gamma}{1+\lambda} r) = F(\frac{1+\gamma}{1+\lambda} v_s + \frac{\lambda-\gamma}{1+\lambda} r)$. Since F is increasing in its argument, it follows that the probability of a sale not occurring – a proxy for time on the market – is increasing in r if $\lambda > 0$ and $\gamma = 0$ (i.e. the bias affects sellers but not buyers), decreasing in r if $\lambda = 0$ and $\gamma > 0$ (i.e. the bias affects buyers but not sellers), and ambiguous if both $\lambda > 0$ and $\gamma > 0$.

In fact, sellers face a trade-off between the price they can get for their home and the amount of time it takes to secure that price. If left-digit bias enters the process only by distorting the seller's valuation (i.e. by increasing their reservation price), then we would expect this to yield higher prices but also longer time-on-market. If, on the other hand, it only increases the buyer's willingness-to-pay, we would expect higher prices and shorter time-on-market. If the truth lies somewhere in the middle, we would expect higher selling prices but ambiguous effects on time-on-market. Thus, the second prediction is: *the effect of r on the probability of a sale not happening is positive if sellers are inattentive, negative if buyers are inattentive, and ambiguous if both are inattentive.*

3. Background and data

3.1. The UK housing market

In the UK, most houses are listed under a sole agency agreement (Merlo and Ortalo-Magne, 2004). This means that a single real estate agency is in charge of all activities from the time when the property is listed until it is either sold or withdrawn. This agent is hired by the seller. This is different from some countries where two real estate agents, one for the buyer and the other for the seller, are involved in a transaction. The agent collects information about the property, determines list prices with the seller, and advertises the property on their website, in brochures, and on property portal websites.

Agents then arrange viewings for potential buyers and ask them to make an offer if they are interested. The buyer who bids the highest normally gets the property. But sellers also choose buyers depending on other conditions, such as whether they are paying in cash or need a mortgage, whether they are in a chain or not, and — if so — the length of the chain. The list price usually shows the approximate amount that the seller hopes to achieve. However, it is not the seller's reservation price, or the lowest acceptable price. Final sales prices can be higher or lower than list prices, depending largely on the quality of the property as well as market conditions.

The price history of properties has become increasingly accessible in the UK in recent years. HM Land Registry (LR hereafter) provided price-paid data as a commercial service between 2004 and 2012 and as a free service since 2013. At first, the service was inconvenient for tracking one specific property. Computational skills were required because the data were expressed in a large comma-separated value file. Later, LR introduced a search portal where previous prices can be easily found by postcode (https://www.gov.uk/search-house-prices). The largest two property portals, Rightmove and Zoopla founded in 2000 and 2008, respectively, provide the price paid data in a more user-friendly manner through a simple search with addresses, immensely improving public data accessibility. As a result, potential buyers are increasingly aware of previous sales prices. This feature makes the UK housing market a perfect setting to test the hypotheses.

3.2. Data

I bring together data from multiple sources to track house transactions, observe property characteristics, and determine the characteristics of the buyers and sellers. The primary data source is LR, which records almost the universe of residential property sales in England and Wales sold at full market value and submitted for registration[2]. I use the house transaction data of the Greater London region from 1995 to 2017. For each sale, the dataset contains transaction prices, transaction dates, and precise postcodes and addresses (street name, street number, and apartment number when the property belongs to a multi-unit building). The dataset records three property attributes, namely, the tenure type (leasehold or freehold)[3], the property type (flat, terraced, semi-detached, detached, or other), and whether or not the property is a new build. The analysis of transaction prices presented in this research relies on the identification of repeat sales of the same property. I consider that two sales occur on the same property when they share the same postcode, flat number (for flats), street name, and street number.

I apply some restrictions to the sample. Following previous literature, such as Levitt and Syverson (2008) and Ben-David (2011), I remove transactions with extreme prices, which are defined as the top and bottom 1 percent of prices. I focus on owner-occupied and non-investment units as much as possible because some investors buy houses and quickly resell them for a profit after repairs and updates to the houses, which is commonly referred to as house "flipping". In the absence of perfect measures of house attributes, those repeat sales will lead to misleading results. I thus exclude properties where the seller owned the property for less than six months. For the graphical and parametric analyses, I further limit the sample to those with previous prices at or under £500,000 because few charm prices are used when property values exceed £500,000. The threshold results in the loss of a relatively small number of properties, as 95.9 percent of the whole sample remains. After imposing these restrictions, the remaining sample contains 2,919,584 transactions for 1,736,326 properties. A total of 1,144,572 sales find at least one matched purchase in the sample[4].

Considering that LR provides a limited number of data fields, I use two other datasets to supplement LR data. First, I construct a dataset of property attributes by extracting keywords in the scraped advertisement texts from Rightmove (RM, hereafter). RM is the largest online property listing website in the UK. It lists all properties for rent or for sale via estate agents· Under the section 'Price Paid Data', RM provides property attributes that are previously listed in rental/sales advertisements on this website. All listed properties contain information on the number of bedrooms. I obtain other attributes through textual analysis. In particular, I search in the texts to identify whether they mention 'parking', 'garden' and 'patio'. Sellers or estate agents are presumed to specify those characteristics in the advertisement when the property has those selling points. I merge RM property attributes with the LR dataset by postcode and exact address including apartment number. Precise matching is possible via primary addressable object name (PAON), which records the house number or property name, and secondary addressable object name (SAON), which records apartment number when a property has been divided into separate units. Some 362,267 repeat sales, i.e. 31.65% of the LR sample, remain after the match.

Table 1 shows the descriptive statistics of the LR and LR – RM matched samples in column (1) and (2), respectively. I distinguish the number of sales and the number of properties in the first two rows to highlight cases of repeat sales. The number of sales is constantly lower

[2] Details on the excluded data can be found at https://www.gov.uk/guidance/about-the-price-paid-data#data-excluded-from-price-paid-data.

[3] In the UK, two fundamentally different forms of legal home ownership, namely, freehold and leasehold, exist. A freeholder owns a property and the land it stands on outright, in perpetuity, whereas a leasehold owner only leases from the freeholder to use the property for a limited number of years, which is usually 99, 125 or 999. For a detailed review, see Giglio, Maggiori and Stroebel (2015) and Bracke, Pinchbeck and Wyatt (2018).

[4] I identify repeat sales and previous sales prices before applying sample restrictions so that a transaction will only be matched to the closest prior.

Table 1
Descriptive statistics.

—	(1) Sales with previous purchase in 1995–2017 (LR)		(2) Sales with previous purchase in 1995–2017 (LR – RM matched)		(3) Local linear sample		(4) Zoopla sample, 2014-2019	
Sales (*N*)	1,144,572		362,267		311,181		43,220	
Properties	748,872		222,123		202,114			
—	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Current sales price (£)	295,164	192,103	319,810	194,381	321,136	194,885	457,034	226,391
Previous sales price (£)	175,258	104,273	184,498	103,327	188,999	103,224	221,485	100,692
Holding period (year)	5.88	4.07	6.53	4.35	6.25	4.39	10.52	4.78
Freehold	0.43	0.49	0.56	0.5	0.57	0.50	0.66	0.47
Property type								
Flat	0.56	0.50	0.42	0.49	0.38	0.48	0.41	0.49
Terraced	0.28	0.45	0.37	0.48	0.42	0.49	0.33	0.47
Semide-tached	0.13	0.33	0.17	0.37	0.17	0.37	0.21	0.41
Detached	0.03	0.17	0.04	0.19	0.04	0.19	0.05	0.22
Other	0	0.04	0	0.03	0.00	0.03	0.00	0.03
Parking			0.2	0.4	0.19	0.39	0.23	0.42
Garden			0.43	0.49	0.42	0.49	0.43	0.50
Patio			0.02	0.13	0.02	0.13	0.01	0.12
No of bedrooms			2.52	1.03	2.55	1.04	2.67	1.03
Time on the market (day)							292.68	303.21
Total page views							1,066.86	1,146.24

Notes: Column (1) is repeat sales identified in the HM land registry. Column (2) is the main dataset used for the graphical analysis and the regression discontinuity approach. Column (3) is the local linear sample used for the repeat sales approach. It includes sales that have previous prices within a distance of £1,000 from round number thresholds. Column (4) is a subsample of the local linear sample for which listing information is available from Zoopla during 2014 – 2019. Time on the market is defined as days from when an advertisement is posted online to when it is removed.

than the number of properties. In general, repeat sales prices are higher than previous prices (approximately 1.6 times for mean values), indicating that nominal house prices are growing over time. The average gap between two consecutive sales for one property is six years. In order for a transaction to be in the matched sample, a property has to be listed for sale at least once after RM was founded in 2000. Thus, the matched sample is inevitably skewed towards recent transactions. This is reflected in the mean sales price in Table 1 – the matched sample has higher sales prices than the LR sample. But the high price does not present problems for this research because charm-round pairs are evenly distributed in all price ranges. Flats and leaseholds are under-represented in the merged sample (42% versus 56% for flats, and 44% versus 57% for leaseholds). The main reason is that flats, most of which are also leaseholds, were sold directly by developers or estate agents rather than listed at property portals when they were new-build properties. The chance that they are included in RM's records is therefore comparatively low.

Second, I estimate an LTV ratio at a postcode level on a quarterly basis to gauge financial constraints. Equity-constrained sellers tend to set high prices to enable them to fund the down payment of their next home, and they usually obtain high prices (Anenberg, 2011; Genesove and Mayer, 1997; Stein, 1995). A standard gauge of the equity constraint is the LTV ratio. An estimate is obtained with the following data. UK Finance, which represents the banking and finance industry, provides the total amount of borrowing outstanding (2013Q2–2018Q1). I adjust the amount by average household size (2.4) from the 2011 Census data, homeownership rate (64 percent) from the Office for National Statistics and mortgage buyer rate (75.9 percent) from the LR. The LTV ratio is obtained by dividing the outstanding loan level that every homeowner with mortgages holds by the average home value. I derive the average home value from the average sales price by postcode sectors and quarters in the transactions from LR. Appendix A provides additional details about the data and procedures.

4. Graphical analysis

4.1. Bunching at round and charm numbers

Fig. 1(a) presents the distribution of house prices below £500,000 in the UK from 1995 to 2017, and Fig. 1(b) zooms in on the range of £300,000-£400,000 to provide a clearer view. Among the different levels of roundness (multiples of £10, £50, £100, £500, £1,000, etc.), bunches are the most pronounced at each £5,000 multiple, accounting for 52.5 percent of all transactions[5]. A finer unit of £1,000 is less common but still pronounced. Fig. 2(a) shows the distribution of the right-most three digits in house prices. A total of 78.7 percent of the prices end with 000, that is, multiples of £1,000. Fig. 2(a) also demonstrates the prevalence of charm numbers that are just below round numbers. They have the right-most three digits of *9XX*; some most common examples are 950, 995, 999, 900 and 990. They account for 6.76 percent of all prices and 32.2 percent of the non-£1000-multiples. The two price

[5] The remarkable surge at £250,000 in Figure 1(a) is caused by the slab property tax system and the cut-off point of £250,000 where the SDLT rose from 1 percent to a much higher percentage. So are the exceptions around tax rate thresholds of £250,000 and £500,000 in Figure 2(b), where prices bunch under the tax threshold. I discuss the tax influences in section 7.1.

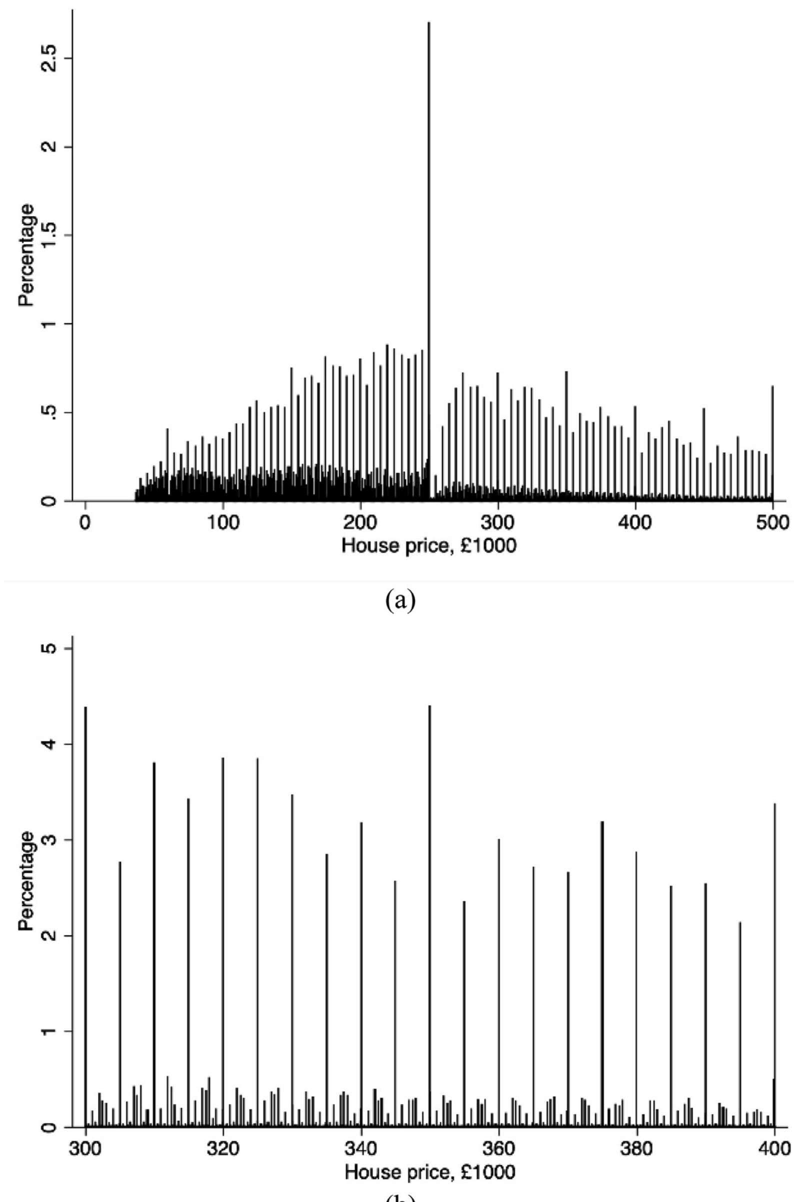


Fig. 1. Distribution of House Prices in the UK (1995–2017)
Notes: (a) and (b) shows the distribution of house sales prices. Panel A plots a histogram of the frequency (in percentage points) of sales occurring at each given price, in bins of size of £10. Panel B reports the same histogram for the interval of prices £300,000-£400,000.

formats account for more than 80 percent of all house prices. Fig. 2(b) presents the proportion of round prices and charm prices under different price magnitudes. The use of charm prices decreases and that of round prices increases with the increase in prices.

4.2. Comparison of charm and round prices

I begin the analysis with plots of previous and current prices in Fig. 3. I divide the previous prices into bins of equal sizes of £1,000, and prices in the same bin have the same left-most three digits (two digits for five-digit numbers), such as: bin [160,000, 161,000) with right-most digits of 160; and bin [169,000,170,000) with right-most digits of 169. Panel (a) in Fig. 3 plots the average current price versus the left-most three digits in the previous price, ranging from £100,000 to £200,000. The dot sizes are proportional to the number of transactions in each price bin. An extended version for prices between £35,000 to £500,000 is presented in Fig. B.1 in the appendix. A pattern that emerges is that the relationship is neither linear nor continuous. Discontinuity occurs when the number at the ten thousands' column changes from 4 to 5 or from 9 to 0. For instance, the average subsequent price increases from £256,277 to £262,638 by £6,361 from bin [164,000, 165,000) to bin [165,000, 166,000), whereas the average subsequent price increases to £263,280 by £642 from [165,000, 166,000) to [166,000, 167,000). Similarly, the average sales prices significantly increase from bin [179,000, 180,000) to bin [180,000, 181,000), that is, £274,702 to £284,721. Observation numbers, as indicated by dot sizes, are considerable, with an average of 4,342 and a minimum of 1,352. In particular, charm number and round number bins are the most populated. Hence, the discontinuity is unlikely to be misled by a few dominant transactions. Notably, the pattern corresponds to the discontinuity predicted by the theoretical model, suggesting that people's perceptions of previous prices may have an abrupt increase when a change occurs in the 10,000's column.

To probe the discontinuity further, I focus on the small area where the number is slightly lower or higher than a £10,000 multiple. This condition applies when the number in the ten thousands' column changes by one unit, for example, from £179,995 to £180,000. The first group of prices are round prices, which are defined as multiples of £10,000. The second group are charm prices, which are defined as prices £1–£100 under those multiples. I calculate the average of the subsequent

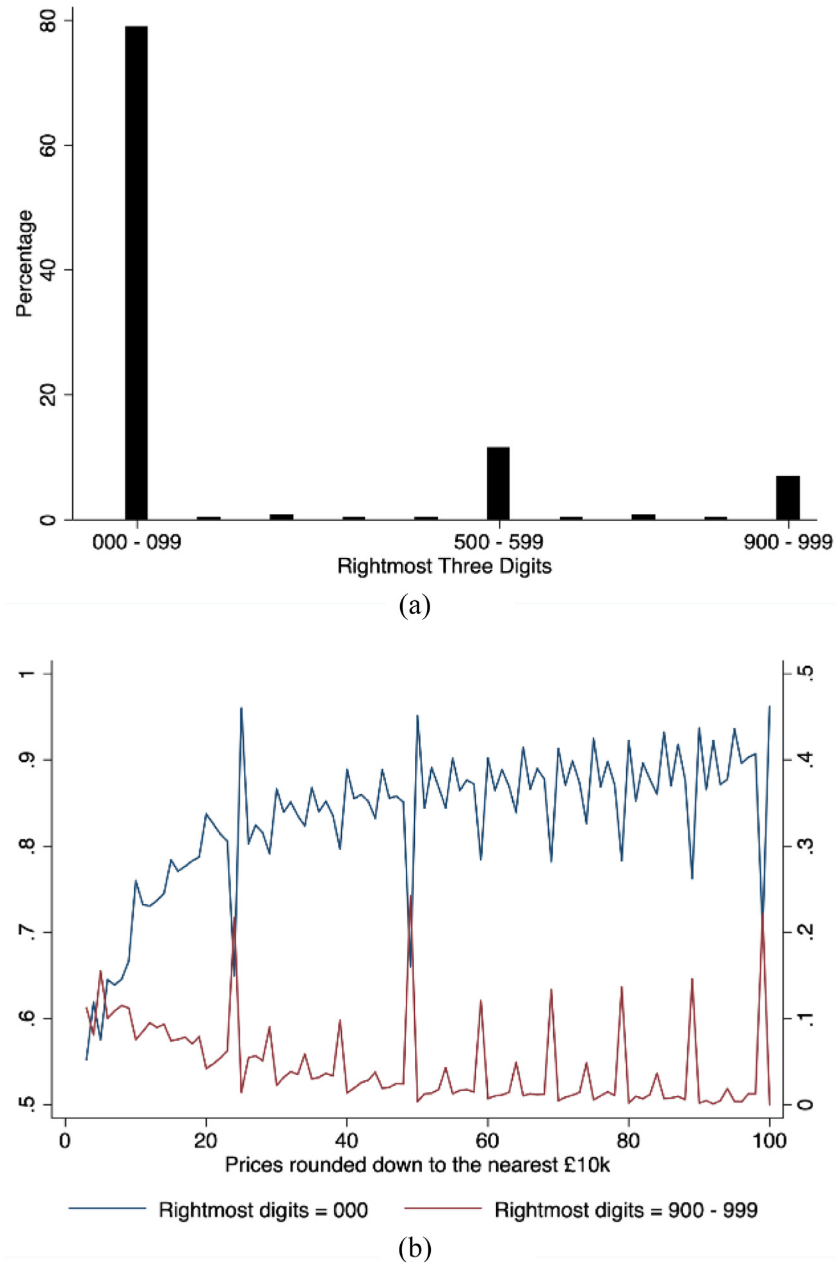


Fig. 2. Bunching at Charm and Round numbers
Notes: (a) shows the distribution of the right-most three digits in house sales prices. (b) shows the percentage of prices being round numbers (right-most three digits = 000) and charm numbers (right-most three digits ≥ 900) varying with property values.

house prices for each group and plot it for the charm (dot) and round groups (cross) in panel (b) of Fig. 3. Each round price and its neighbouring charm prices are presented as pairs to facilitate direct comparisons. The method essentially resembles regression discontinuity design at thresholds of each £10,000 multiple. At each threshold, the treatment is a higher number in the ten thousands' column. Hence, previous prices slightly lower than the threshold are in the control group whereas previous prices slightly higher than the threshold are in the treatment group. I use a small bandwidth of £100. For instance, the round price of £234,000 and the charm prices in (£233,900, £234,000) form a pair and are placed at the same level on the x-axis. Crosses are constantly higher than dots in Fig. 3(b), indicating that round groups have consistently higher subsequent prices than charm groups.

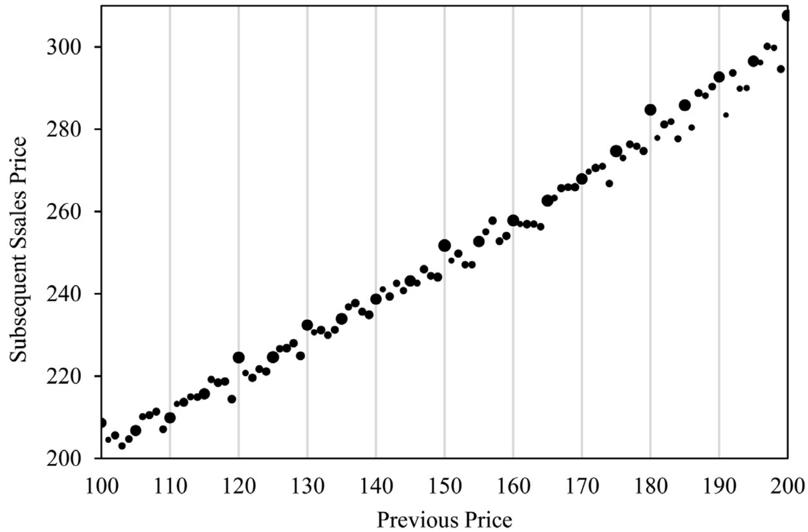
5. Parametric analysis

In this section, I develop an empirical strategy to evaluate whether homes sold at previous prices with different formats have different subsequent prices. As the first step, I use a regression discontinuity design and explore the discontinuous effect of the previous purchase price on the subsequent price. I set a baseline model motivated by the repeat sales hedonic model to quantify the average discontinuity and avoid the omitted variable bias. Then, I augment the baseline specification and the sample to provide estimates at different roundness levels of numbers to eliminate price distortions caused by transfer taxes and to explore the heterogeneity in the discontinuity across properties and over time.

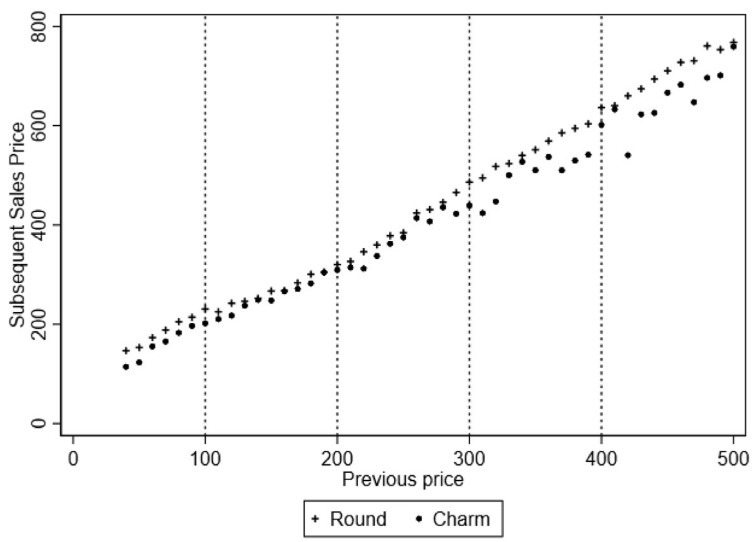
5.1. Discontinuities estimates

In the spirit of the regression discontinuity design (Athey and Imbens, 2017; Lee and Lemieux, 2010), I first utilise the following specification to empirically evaluate the existence of discontinuities at round number previous prices:

$$\ln P_{ijst} = \alpha_0 + f\left(Ln\ P_{is}\right) + \sum_{k=3}^{50} \alpha_k D\left[P_{is} \geq j \times\ 10,000\right] +\ \boldsymbol{X}_i'\beta + \delta_t + \varphi_j + e_{it}, \quad (1)$$



(a)



(b)

Fig. 3. Graphical Presentations of Discontinuity
Notes: Previous prices (x-axis) and subsequent sales price (y-axis) are in 1,000 units. (a) illustrates the average subsequent price against each previous price bin of a £1000 bin width, and the dot sizes are proportional to the number of properties in the bins. (b) compares the average price for properties with round previous prices and charm previous prices at each £10,000 multiple in pairs.

The dependent variable is the natural log of the current price for house i in district j previously transacted at time s and currently at time t. Function $f(\ln P_{is})$ is a flexible function of the log of the previous price to capture the smooth patterns where current house prices anchor to previous house prices. The variables of interests are the set of indicator variables D(·) to determine whether the previous price crosses a round-number threshold, that is, a £10,000-multiple in this specification. The current price is predicted to have a discontinuous jump when the previous price crosses a £10,000-multiple. The parameters (α_3, α_4, $\ldots, \alpha_{50}$) that capture the discontinuous jumps at various multiples are thus expected to be positive. Following the hedonic pricing model (Rosen, 1974), I include a vector of property-specific characteristics, $\boldsymbol{X}_i'$, that are described in Table 2, that is, dummy variables to indicate whether there is a garden, a patio, or a school located nearby, whether the property is a freehold or a leasehold, dummy variables for the number of bedrooms, property types and locational attributes. δ_t defines a time fixed effect for the year of transaction. φ_j defines the district fixed effect to absorb the locational effects and unobserved time-invariant property characteristics that are homogeneous to properties within the same district. α is a constant, and e_{it} is the error term.

Fig. 4 shows the regression estimates for each £10,000-multiple. I include a fifth-order polynomial for function $f(P_{is})$. The discontinuity level is highest when previous prices are below £100,000, stays positive and slowly increases from £100,000 to £250,000. This level then encompasses huge variations and decreased precisions as indicated by wide confidence intervals[6]. The overall pattern features a U-shape. In addition, I test the robustness of the results to include holding time — defined as the days between two sales — as an additional control, and two-way fixed effects of sales year and month to control for seasonality. Results are similar to Fig. 4 and are presented in appendix Fig. B.2.

5.2. Repeat sales approach

A typical challenge in the hedonic price model and the specification in Eq. (1) is that the potential bias that is related to the omitted variables confounds with the variable of interest. I utilise an augmented repeat sales approach to sidestep this issue (Agarwal et al., 2015; Ben-David, 2011). In particular, the dependent variable is the difference in

[6] The abrupt jump at £260,000 is a result of price distortions caused by the UK land tax scheme, which will be covered in detail in section 7.1.

Table 2
Baseline effect size estimates.

	(1)	(2)	(3)	(4)	(5)	(6)
Round	0.006***	0.053***	0.040***	0.041***	0.043***	0.044***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Property characteristics			√	√	√	√
Fixed effects						
Sale year		√	√	√		
District			√	√		
Magnitude				√	√	√
District × sale year					√	√
Sale year × month						√
Observations	311,181	311,181	311,181	311,181	311,181	311,181
R-Squared	0.000	0.721	0.740	0.793	0.810	0.818

Notes: This table provides the estimates based on Equations (2) on the full sample. All the regressions are OLS estimates and standard errors are clustered at the property level. The dependent variable is the log of the between the current house price and the prior house price. The independent variables include the dummy variable *Round* which equals to one if the previous purchase price is a multiple of 1,000, property characteristics, district fixed effects, previous and current year fixed effects, magnitude of previous price fixed effects and two-way fixed effects of districts and sales year (previous and current). Standard errors are shown in parentheses.
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

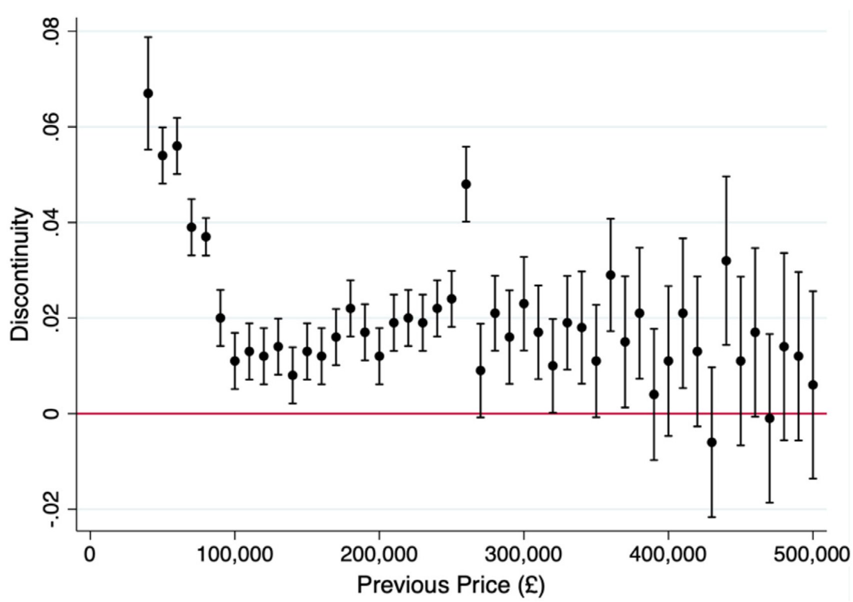


Fig. 4. Estimates from Regression Discontinuity Design
Notes: This figure provides estimates based on Eq. (1). All the regressions are OLS estimates, and standard errors are clustered at the property level. The dependent variable is the log of the house price. The independent variables include a flexible function of the log of the previous price, a set of dummy variables for whether the previous price crosses a round number threshold, property characteristics, district fixed effects, year fixed effects. Standard errors are shown in parentheses.
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

house prices between two consecutive transactions on the same property[7], thereby removing the effects of time-invariant property-specific characteristics, including unobserved ones. Therefore, the price difference for property i with the previous transaction at time s and the latter at time t is represented in Eq. (2).

$$\ln P_{it} - \ln P_{is} = \ln(P_{it}/P_{is}) = \alpha_0 + \alpha_1 Round_i + \boldsymbol{X}_i'\beta + \delta_s + \delta_t + \varphi_j + e_{ist}, \quad (2)$$

where α_0, $\boldsymbol{X}_i'$, δ_t and φ_j are the same as defined in Eq. (1). δ_s represents an added set of dummy variables to denote the year of previous transaction. e_{ist} is the new error term. δ_s and δ_t control for the temporal variation in average house price levels in the Greater London region and absorb the effect of the holding period (in years)[8]. A new variable *Round* defines whether the previous price is a round number (=1) or a charm number (=0). In the neighbourhood of a round number, an increase in one of the left-most digit, such as from £239,950 to £240,000, renders the variable *Round* from zero to one. The charm format in the previous price is predicted to bias the reference point downwards, and the subsequent transaction price will be lower compared with its round price counterpart. Hence, the main variable of interest, *Round*, is expected to have a positive coefficient.

To estimate Eq. (2), I restrict the sample to transactions with previous prices that are either a round price (multiples of £1,000) or a charm price (no more than £100 below a round price). Comparing the sales outcome in the two groups is similar to “the penny-wise and pound-foolish” experiment that researchers utilise to evaluate a consumer’s response to charm pricings (Thomas and Morwitz, 2005). The remaining sample contains 311,181 repeat sales, in which 277,629 are previously transacted at round prices and 33,552 at charm prices. Descriptive statistics for the local linear sample are given in the third column in Table 1. Both the prices and the observable characteristics are comparable to the full sample in column (2).

[7] Multiple resales are treated as independent observations.

[8] Year fixed effects are coarse controls as the holding period is accounted for in years. For robustness, I include holding time (in days) as an additional control. The conclusion still holds. Table B.2 in the appendix presents the results.

This research does not use the regression discontinuity design. Therefore, neither continuity of the potential outcomes nor local randomisation is required. Instead, I include housing attributes in the regression to control for potential differences. However, it is interesting to see if properties with charm price differ from those with round prices in terms of observables. I conducted balance checks by taking housing attributes as the dependent variables and run Eq. (2) without controls and fixed effects. The results are presented in Table B.1 in the appendix. It shows that the two groups are not different in terms of whether it has a garden, whether it has a patio, or whether it has a school nearby; nor are they different in terms of tenure type (freehold or leasehold) or building type (flat or house). But they differ in the number of bedrooms. This difference highlights that including this variable in the regression is essential.

Table 2 presents the estimation output. Standard errors are clustered by property to account for potential correlation across sale observations for the same property. Column (1) includes only the variable *Round* without control variables. The coefficient, 0.006, is positive as expected. This coefficient implies that the number in the thousands' column increasing by one unit is associated with higher prices in the subsequent transaction. Columns (2) to (3) add control variables and restrictive fixed effects to the model. Column (2) adds the fixed effects for the years when the previous and current sales take place. Column (3) adds property attributes and district fixed effects, as specified in the full specification of Eq. (2). The remaining effect size is 0.04, indicating that the subsequent sales prices are 4 percent higher when previous prices are one unit higher in the thousands' column, everything else being equal.

Round prices are related to high previous prices and high price changes. However, it is difficult to disentangle whether the driving factor is the increasing price magnitude or the digit change effect. Therefore, in column (4) I include price magnitude fixed effects. The implementation is done by pairing charm and round prices with the same magnitude when their values have less than £1,000 difference. For instance, £239,900 and £240,000 have the same magnitude, and £249,900 and £250,000 have the same magnitude. In essence, similar to the pairing comparisons in the graphical analysis, this specification enables us to compare prices around one round number. The coefficient for *Round* captures the average treatment effect size. It remains significant and positive. The effect sizes have similar magnitudes as in column (3), that is, a one-unit change in the thousands' column of the previous price is associated with a 4 percent price premium in the subsequent transaction price.

For robustness, column (5) and (6) presents two alternative specifications. In column (5) I replace the separate year and district fixed effects with their interactions, that is, $\delta_s \times \varphi_j$ and $\delta_t \times \varphi_j$. Together, the two terms capture the average district-level changes in the price increase for each combination of the previous-current transactions investigated. Another benefit for the two-way fixed effects is to absorb much of the variation related to financial constraints (Ben-David, 2011; Mian and Sufi, 2009). In column (6) I add the sales year × month effect to account for potential seasonality issues in house prices. Coefficients for *Round* are not affected. Additional robustness checks can be found in Table B.2 and Table B.3 in the appendix. Specifically, I add holding time (in days) as an additional control in Table B.2, and I allow heterogeneous slopes on either side of round number thresholds in Table B.3. They do not change the main conclusions.

5.3. Roundness at multiple levels

The theoretical model predicts that more rounded numbers can have a stronger effect. For example, £300,000 is more rounded than £230,000. The digit change from £299,950 to £300,000 occurs in the hundred thousands' column, whereas that from £229,950 to £230,000 takes place at the ten thousands' column. Varied columns in a multi-digit number have different column values, attract differentiated levels of attention, and may have differentiated effect size. The hypothesis from section 3.1 is that people read from left to right and are more attentive to left-most digits. Hence, the more rounded the numbers are, the higher the effect size will be.

To test this prediction, I utilise an augmented specification from Eq. (3) as follows:

$$\ln(P_{it}/P_{is}) = \alpha_0 + \alpha_1 Round_i + \sum_{k=2}^{5} \rho_k Round_i \times Multiple_k + \sum_{k=2}^{5} \eta_k Multiple_k + X_i'\beta + \delta_s + \delta_t + \varphi_j + e_i \quad (3)$$

where $Multiple_i$ ($i = 1, 2, \ldots, 5$) represents a set of dummy variables that corresponds to the level of roundness. Let $Multiple$ denote five levels of roundness. Level 1 to 5 correspond to multiples of £1,000, £5,000, £10,000, £50,000, and £100,000, respectively. For example, £234,000 is a multiple of £1,000, and $Multiple$ equals 1. *Multiple* indicates the biggest factor and the highest level of roundness when a round number is the multiple of more than one aforementioned factor. For example, £300,000 is a multiple of £1,000, £5,000, £10,000, £50,000, and £100,000, where £100,000 is the largest factor, and *Multiple* equals 5. $Multiple_i$ ($i = 1, 2, \ldots, 5$) equals to one when $Multiple_i = i$ and zero otherwise.

In Eq. (3), I include $Multiple_i$ and its interaction term with $Round$. The main coefficients of interest are α_1 and ρ_k. I expect positive estimates based on the hypothesis. The effect of being at a £1,000 multiple is α_1, the effect of being at a £5,000 multiple is $\alpha_1 + \rho_2$, …, and the effect of being at a £100,000 multiple is $\alpha_1 + \rho_5$. All other variables are the same as in the previous step.

Column (1) in Table 3 presents the multilevel estimates for the full sample. As expected, the coefficients for the interaction terms of £5,000, £10,000 and £100,000 are all positive and are increasing. The coefficient for *Round* × £5,000-multiple is not significant, indicating that the effect sizes at £1,000-multiples and £5,000-multiples are not sig-

Table 3
Effect size at multiple levels.

	(1)	(2)
Round	0.043***	0.047***
	(0.002)	(0.002)
Round × £5,000-multiple	0.003	0.003
	(0.003)	(0.003)
Round × £10,000-multiple	0.003	0.003
	(0.003)	(0.002)
Round × £50,000-multiple	−0.015***	-0.016***
	(0.003)	(0.003)
Round × £100,000-multiple	0.012***	0.013**
	(0.004)	0.003
Property characteristics	√	√
Fixed effects		
Magnitude	√	√
District × sale year	√	√
Sale year × month		√
Observations	311,181	311,181
R-Squared	0.810	0.818

Notes: This table provides the estimates based on Eq. (3) on the full sample. All the regressions are OLS estimates, and standard errors are clustered at the property level. The dependent variable is the log of the difference between the current house price and the prior house price. The independent variables include the dummy variable *Round* which equals to one if the previous purchase price is a multiple of 1,000, property characteristics, district fixed effects, previous and current year fixed effects, magnitude of previous price fixed effects and two-way fixed effects of districts and sales year (previous and current). Standard errors are shown in parentheses.
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 4
Heterogeneity by transaction year.

	Coefficient
Round	0.081***
	(0.005)
Round × Year (2000 – 2007)	-0.032***
	(0.005)
Round × Year (2008 - 2009)	-0.040***
	(0.006)
Round × Year (2010 onwards)	-0.040***
	(0.005)
Property characteristics	√
Fixed effects	
Magnitude	√
District × sale year	√
Sale year × month	√
Observations	293,722
R-Squared	0.820

Notes: This table provides the estimates based on Eq. (2). All the regressions are OLS estimates, and standard errors are clustered at the property level. The dependent variable is the log of the difference between the current house price and the prior house price. The independent variables include the dummy variable *Round* which equals to one if the previous purchase price is a multiple of £1,000, interaction terms of *Round* and three time period dummies, property characteristics, district fixed effects, previous and current year fixed effects, magnitude of previous price fixed effects and two-way fixed effects of districts and sales year (previous and current). Standard errors are shown in parentheses.
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

nificantly different. Notably, the coefficient for the £50,000-multiple interaction term is negative, possibly because of the prominent distortions caused by the £250,000 threshold. Column (2) adds the two-way fixed effects of sales year and month to control for seasonality, and the conclusions remain unchanged.

5.4. Heterogeneity

The estimated overall effects in the full sample mask economically interesting heterogeneity. The effect size may vary with time because of changing market settings – economic shocks, for example – and the introduction of the Internet. Price information becomes transparent with the omnipresence of the Internet. In this section, I examine the effect size differences in four time periods, namely, before 2000, 2000–2007, 2008–2009 and after 2009. The cut-offs are selected by price information accessibility. Before 2000, previous price information is not publicly available to home buyers. Real estate agents may know the price from sellers, but they may not give an accurate number and tend to round up. The period 2000–2008 is when online information services were developing and the exact price became increasingly visible. Since 2008, real estate information has become easily accessible on the Internet. I set 2008–2009 as a separate period to represent the financial crisis when household behaviours may exhibit differential patterns.

I pool observations and include interaction terms with *Round* for each time period and re-estimate the effect size with Eq. (2). Table 4 reports the results. Effect sizes for pre-2000, 2000 – 2007, 2008 – 2009, and 2010 onwards are all statistically significant and monotonically decreasing: 8.1 percent, 4.9 percent, 4.1 percent, and 4.1 percent. Two insights emerge from these results. First, the strong effect for pre-2000 suggests that some of this bias is coming from the seller side. Suppose that this effect was driven entirely by buyers being affected by previous prices. The effect before 2000 should be smaller, or non-existent, because price information was not publicly available then. Second, a plausible explanation for the decreasing trend is the advent of the Internet and the resultant improved channel of information dissemination. Price information has become visible to buyers since pre-2000. Salience of information attenuates inattention (DellaVigna, 2009). When property

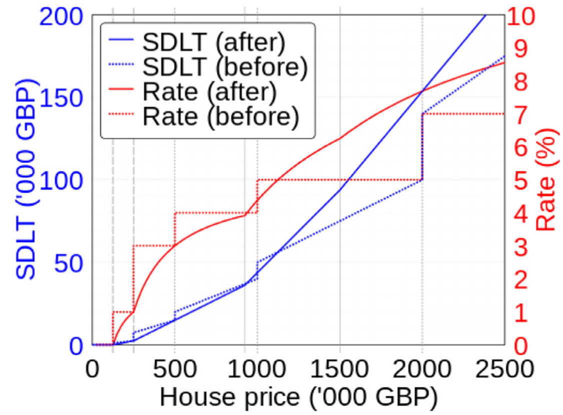


Fig. 5. SDLT for Individuals Before and After 4 December 2014
Source: https://en.wikipedia.org/wiki/File:UK_stamp_duty_2014_change.svg

portals display prices clearly, people absorb price information better, and hence the effect of left digit bias reduced.

6. Alternative explanations

6.1. Tax considerations

The SDLT in the UK is paid by the buyers and is a percentage share of the purchase price of a property. Before 4 December 2014, it is a progressive 'slab' tax for the entire purchase, which increases at certain thresholds (Fig. 5). For example, at the second threshold, that is, £250,000, the stamp duty payable is £2,500 (tax rate = 1 percent) whereas the tax rate becomes 3 percent and the tax increases to £7,500 when the purchase price exceeds £250,000. The 'slab' tax scheme is believed to have distorted house prices and led them to bunch before tax rate thresholds (Best and Kleven, 2018). In particular, houses valued close to the tax thresholds are normally transacted at prices at or below the tax threshold to avoid a substantial tax increase, so creating a trading "hole" above the tax threshold. Thus, previous prices at or under tax thresholds are likely to be lower than the house value. Under-pricing components will be corrected in subsequent transactions when house values are no longer around any tax threshold. Under-pricing affects the results, yet in an arbitrary way. Bunching at round prices (the threshold) will bias the discontinuity and coefficient estimate for *Round* upward, whereas bunching at charm prices (under the threshold) will bias the discontinuity and coefficient estimate for *Round* downward. The magnitude can be measured and interpreted at two dimensions: either an extensive margin (more transactions) or an intensive margin (more under-pricing).

In the sample, more than 99 percent of the previous transactions occur before 4 December 2014. The prices at or below tax thresholds in the sample are not reflective of the house value, and including them in the empirical analysis will bias the estimate. To mitigate the concerns of tax-related distortions and to reaffirm the identification of the discontinuity, I utilise the doughnut approach developed by Hilber and Lyytikainen (2017) by dropping transactions when their previous purchase prices are around the tax thresholds. Tax thresholds have undergone frequent changes over time. In the examined time period and price range, there are six thresholds: £60,000, £120,000, £125,000, £175,000, £250,000, and £500,000. They are effective at different times (see Table B.4 in the appendix for details). I exclude observations with prices around tax thresholds that are in effect at the time of transaction.

Columns (1) and (2) in Table 5 present the average effect size for the tax doughnut sample with two sets of control variables. The effect is significant in both models. As expected, tax-related price distortions affect estimates of the effect. The variable *Round* has slightly larger coefficients (0.044 and 0.046) than the full sample estimates (0.041 and 0.43).

Table 5
Effect size in the doughnut sample.

	(1)	(2)	(3)	(4)
Round	0.046***	0.047***	0.043***	0.043***
	(0.001)	(0.001)	(0.002)	(0.002)
Round × 5k-multiple			0.002	0.003
			(0.003)	(0.003)
Round × 10k-multiple			0.001	0.001
			(0.003)	(0.003)
Round × 50k-multiple			0.012*	0.012*
			(0.005)	(0.005)
Round × 100k-multiple			0.025***	0.026***
			(0.005)	0.003
Property characteristics	√	√	√	√
Fixed effects				
Magnitude	√	√	√	√
District × sale year	√	√	√	√
Sale year × month		√		√
Observations	292,766	292,766	292,766	292,766
R-Squared	0.795	0.814	0.814	0.820

Notes: This table provides the estimates based on Eq. (3) on the doughnut sample that excludes transactions around effective tax thresholds. All the regressions are OLS estimates, and standard errors are clustered at the property level. The dependent variable is the log of the difference between the current house price and the prior house price. The independent variables include the dummy variable Round which equals to one if the previous purchase price is a multiple of 1,000, property characteristics, district fixed effects, previous and current year fixed effects, magnitude of previous price fixed effects and two-way fixed effects of districts and sales year (previous and current). Standard errors are shown in parentheses.
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Columns (3) and (4) in Table 5 run Eq. (3) on the tax doughnut sample. The coefficients for *Round, Round* × £5,000-multiple and *Round* × £10,000 multiple are similar to the estimates in the two previous columns. By contrast, the coefficients for *Round* × £50,000 and *Round* × £100,000 significantly increase. The differences correspond to the fact that tax thresholds are mostly multiples of £50,000 and £100,000. Estimates from the full sample are biased downward by taxes. After excluding tax-affected observations, effect sizes grow with roundness levels, as predicted in section 3.3. It is worth noting that although excluding tax thresholds eliminates distortions in the housing market, substantial number of records are also removed. A consequence is the inflated standard errors for £50,000-multiple and £100,000-multiple interactions terms.

6.2. Equity constraints

Equity constraints play an important role in housing transactions. People use the equity withdrawn from selling their current home to pay the down payment for their next home. Financially-constrained sellers will hold high reservation prices and are unwilling to accept low house prices—sometimes at the cost of postponing moving decisions—and eventually obtain higher house prices compared with non-constrained sellers (Anenberg, 2011; Genesove and Mayer, 1997; Stein, 1995). In this section, I evaluate whether the premium identified in properties with round-number previous prices is driven by differences in equity constraints, that is, whether buyers of round prices in the previous sales are financially constrained, hold high reservations prices, and achieve high resale prices.

To gauge the equity constraint, I use the ratio between the outstanding loan per household and the average home value at postcode-quarter levels from 2013Q2 to 2018Q1. Appendix A provides the details about the data source and the estimation procedure. Fig. 6 compares the LTV estimate with the LTV statistics released by UK Finance, a trade association for the UK banking and financial sector. The statistics are separately reported for first-time buyers and home movers. First-time buyers have higher LTV ratios than home movers, averaging at 0.74 and 0.66, respectively. The LTV estimate is approximately one-third of the official statistics, with a mean value of 0.26. The deviation arises from three main sources. First, the typical LTV refers to the ratio at loan origination. The outstanding loan level is lower than the loan level at origination because substantial repayments have been deducted. As a result, the estimated LTV is diluted and is lower than the typical LTV. Second, the home value represents its current market value rather than the value when the property was bought. Real house prices never drop in the examined time period, which again leads to a downward deviation. Third, home values in mortgage-financed transactions are difficult to distinguish from cash-financed transactions (approximately 25 percent of all transactions) when calculating home values.

Despite the magnitude difference, Fig. 6 demonstrates a clear common trend for the three lines, which illustrates the accuracy and effectiveness of the estimate at an aggregate level. The estimate using the outstanding loan level and current property values is reflective of homeowners' current financial conditions, whereas LTV at origination is only reflective of their financial conditions at the time of transaction. Hence, the current LTV ratio is widely accepted in the extant literature (Anenberg, 2011; Genesove and Mayer, 2001; Genesove and Mayer, 1997). Admittedly, as the estimate is at a postcode-sector level rather than an individual level, it only stands as a coarse measure. However, the estimate remains an effective measure, as supported by Mian and Sufi (2009)'s finding that equity constraints are strongly associated with geographic location, that is, US zip codes in their paper and postcode sector in this paper.

I match each transaction with the LTV estimate of the quarter when the seller purchased the property, which proxies for the equity constraint. The match causes a substantial loss of sample because a property has to have two or more transactions for the time period from 2013Q2 to 2018Q1 in order to be in the sample. Some 8,679 observations remain after the match.

Eq. (4) presents the empirical model for the estimation.

$$\ln(P_{itp}/P_{isp}) = \alpha_0 + \alpha_1 Round_i + \sum_{k=2}^{5} \rho_k Round_i \times Multiple_k + \gamma LTV_{i,p,q(t)} + X_i'\beta + \delta_s + \delta_t + \varphi_j + e_i, \tag{4}$$

where $LTV_{i,p,q(t)}$ is the equity measure for property i in postcode sector p at quarter q, and q is a function of the transaction year t. The existing literature indicates that the equity effect has a threshold, and only high LTV units, that is, constrained households, are sensitive to equity (Genesove and Mayer, 1997). A typical threshold used to define highly leveraged households is 0.8 (Anenberg, 2011; Ben-David, 2011; Genesove and Mayer, 2001). Hence, the equity effect is empirically evaluated with a variable defined as the product of the LTV ratio minus 0.8 and a dummy variable that equals 1 when LTV is higher than 0.8 and 0 otherwise. I use the same method to define $LTV_{i,p,q(t)}$. Given that the LTV estimates are lower than the typical LTV, a threshold of 0.8 is not applicable. I use the upper quantile of the LTV estimate, that is, 0.33, as the threshold.

Table 6 presents the estimation results. Column (1) provides a baseline model that estimates Eq. (4) with the new sample. It does not include LTV variables. Control variables have the same definitions as in previous sections. Their coefficients have expected signs and are not shown in the table. Column (2) adds the LTV estimate. It has a positive coefficient and is significant at 5 percent significance level. The result is consistent with existing findings that financially constrained sellers hold higher reservation prices, ask for higher prices and usually achieve higher final prices. Coefficients for the key variable, *Round*, are positive and significant in both columns (1) and (2). Adding the LTV estimate does not alter the effect size and significance.

An alternative way to define the equity constraint is to use a threshold. Some evidence suggests that only households with high LTV ratios are sensitive to equity constraints (Genesove and Mayer, 1997).

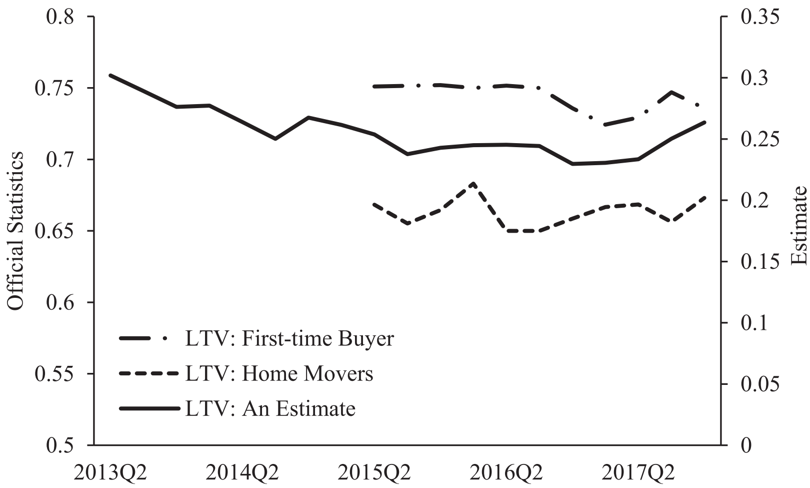


Fig. 6. LTV at Loan Origination versus LTV Estimate
Notes: Official statistics of LTV for first-time buyers and home movers are obtained from UK Finance. They publish aggregate information on mortgages based on the data supplied by members and grossed up to estimate total market size. The left axis shows official statistics, and the right axis shows LTV estimates.

Table 6
Robustness check on equity constraints.

	(1)	(2)	(3)
Round	0.016**	0.016**	0.016**
	(0.006)	(0.006)	(0.006)
LTV estimate		0.063**	
		(0.025)	
High LTV dummy			0.108**
			(0.041)
Property characteristics	√	√	√
Fixed effects	√	√	√
Magnitude	√	√	√
District × sale year	√	√	√
Sale year × month	√	√	√
Observations	8,683	8,683	8,683
R-Squared	0.553	0.553	0.553

Notes: Columns (1)–(3) provides the estimates from Eq. (4). All the regressions are OLS estimates, and standard errors are clustered at the property level. The dependent variable is the log of the difference between the current house price and the prior house price. The independent variables include the dummy variable for the round price, proxy for financial constraints, property characteristics, district fixed effects, previous and current year fixed effects, magnitude of previous price fixed effects and two-way fixed effects of districts and sales year (previous and current). Standard errors are shown in parentheses.
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

A typical threshold used to define highly leveraged households is 0.8 (Anenberg, 2011; Ben-David, 2011; Genesove and Mayer, 2001). Hence, the equity effect is empirically evaluated with a variable defined as the product of the LTV ratio minus 0.8 and a dummy variable that equals 1 when LTV is higher than 0.8 and 0 otherwise. As LTV estimates are lower than the typical LTV, a threshold of 0.8 is not applicable. I use the upper quantile of the LTV estimate, that is, 0.33, as the threshold. The definition method is otherwise the same except for this threshold difference, that is, $(LTV_{isp} - 0.33) * (LTV_{isp} - 0.33)^{+}$. Model (3) presents estimates using the high LTV dummy variable. The size and significance of the variable *Round* is unchanged. Therefore, the effect persists after controlling for sellers' financial constraints.

Notably, the estimated round number premium (1.7 percent) is lower than the baseline estimation in Table 3 (4 percent). This is unsurprising as the sample in this section is biased and the difference reflects heterogeneity of the behavioural bias. First, the sample includes only transactions after 2013, making it similar to the last period sample in section 6.5. The round number premium changes over time. Second, it comprises frequently traded properties, i.e. traded at least twice after 2013. Third, related to the previous point, transactions in the sample involve frequent buyers and sellers who have accumulated knowledge and experience. The knowledge and experience may effectively mitigate their behavioural bias.

6.3. *Pricing errors*

An alternative explanation to the discontinuity identified in t = 0 prices is that charm previous prices may entail a premium. The premium is closely related to the charm pricing strategy that states home sellers attempt to take advantage of buyers' cognitive biases and list prices at charm numbers. The strategy normally leads to a premium in sales prices (Allen and Dare, 2004; Chava and Yao, 2017; Repetto and Solís, 2020). If charm prices at t = –1 are indeed overpriced, the pricing premium disappears at t = 0 and sales prices are lower than their round-priced counterparts. Hence, the discontinuity at t = 0 prices can be attributed to the overpricing component.

This explanation is not plausible for the identified effects for two reasons. First, charm list prices do not necessarily lead to charm sales prices because of price negotiations between buyers and sellers. Second, if properties are intrinsically inferior for the charm group, they should differ in their attributes. The observables in the empirical model are able to capture the differences and therefore the premium is not due to housing attributes. However, as there is always the possibility of missing attributes, for example, whether the property has south-facing windows, I use the following placebo test to probe whether the overpricing component, if any, can explain the identified effect.

Suppose a property has been transacted at three time points: the current transaction at t = 0 and two prior transactions of a closer one at t = –1 and a further one at t = –2. The reference dependence links the previous price at t = –1 and the subsequent sales price at t = 0, such that left-digit bias is present and the left-most few digits of the previous price at t = –1 determine the subsequent sales price. Those left-most digits should not affect the prior transaction price at t = –2 as we do not expect any effect at all. Price differences at t = –2 are therefore a pure reflection of house attribute differences between properties in the charm group and those in the round group without the influence of left-digit bias.

I investigate the relationship of prices at t = –2 and prices at t = –1 following the same steps in the graphical analysis in Fig. 3(a)–(e). In particular, I split the sample into equal-sized bins of £1,000 based on

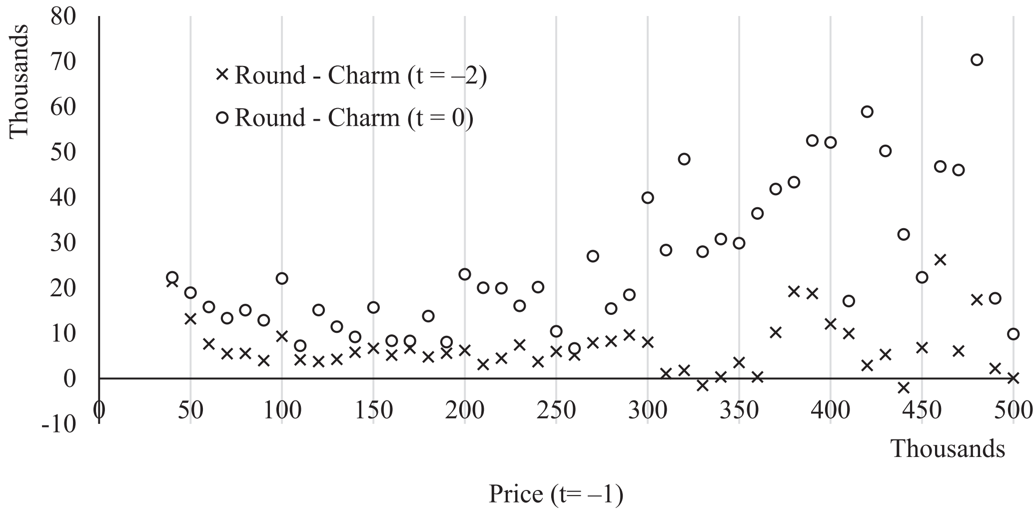


Fig. 7. The Placebo Test

Notes: This figure compares the average current sales price (t = 0) and average sales prices that are prior to the previous transaction (t = –2) against the previous transaction prices at t = –1, in bins of size £1,000. Each dot represents the discontinuity at a round-number threshold. The left axis shows t = 0 prices, and the right axis shows t = –2 prices.

the transaction prices at t = –1 and plot the average transaction prices at t = –2 against the midpoint in each bin. Hence, the sample is restricted to properties that have been transacted at least three times during the investigated period. Fig. 7 compares the discontinuity levels at round numbers at t = –2 (cross) and t = 0 (circle). Circles have constantly higher magnitudes than crosses, with means of £25,455 versus £6,875. Specifically, under £250,000, the discontinuity levels at t = 0 are approximately two times higher than that at t = –1; above £250,000, the discontinuity levels at t = –2 become volatile although they have similar levels as under £250,000. Therefore, graphical evidence supports the coexistence of inattention and pricing errors.

To further eliminate the effects of the observables and market price levels, I conduct similar parametric analysis, as shown in column (7) of Table 3 and substitute the variables at t = 0 with variables at t = –2, such that I can estimate the magnitude of the overpricing component. After controlling for property characteristics, fixed effects for the district, year at t = 0, year at t = –2 and price magnitude, the estimate shows that attribute-related price differences are only 2.6 percent. This is much lower than the 4.4 percent predicted in the baseline regression. Overpricing components alone cannot explain the discontinuity in prices at t = –2. Thus, I confirm the presence of left-digit bias.

7. Who is inattentive?

An interesting question to ask following the empirical findings is who is inattentive— buyers, sellers, or both? A straightforward approach to this question is to investigate the list price. Home sellers for properties previously sold at a round price will set systematically higher list prices than sellers for properties sold at charm prices neighbouring that round price if sellers are inattentive. In this case, price discontinuity is already present when properties are advertised. In later stages of the transaction, buyers may be able to correct the bias if they are rational despite the systematic difference in list prices. However, this inattention-related discontinuity may remain in transaction prices for the three reasons described in the theoretical model — one being a quality signal and the other two being anchoring. In this section, I utilise a sub-sample that has available listing information to identify the inattention of the two counterparties.

I obtain property listing records from 2014 to 2019 from Zoopla, the second largest property listing portal in the UK, and match them with the above dataset by property address[9]. I use this sub-sample to investigate the mechanism behind the price discontinuity. Column (4) in Table 1 provides descriptive statistics for this sub-sample. The most obvious difference between these 43,220 sales and the local linear sample is the much higher prices and longer holding period.

In Fig. 8, I plot prices against previous sales prices using the same method as Fig. 3(b). The left panel presents the effect on sales prices. Round previous prices (plus signs) are associated with higher subsequent sales prices (circles). This pattern is consistent with previous empirical findings. The right panel illustrates the effect on list prices. The discontinuity pattern is again obvious. In other words, sellers set systematically higher list prices when the previous price is a round number than if it is a neighbouring charm number.

I then run Eq. (2) with different dependent variables to empirically test the effect of round prices on subsequent sales outcomes. Row (1) in Table 7 replicates the baseline regression with the new subsample. Model 1 includes the full set of control variables and fixed effects as in column (6) of Table 2, and model 2 includes an additional dummy to indicate whether list prices are charm number or not to account for the charm pricing effect. Results are similar to those in Table 2. Rounded previous sales prices entail a 4 percent sales price premium. Row (2) takes list prices as the dependent variable. The coefficient for *Round* has a positive sign, which confirms the theoretical conjecture that inattention-biased reference points change sellers' reservations prices and lead to systematic differences in list prices. The coefficients in row (1) are only slightly lower in magnitude than those in row (2), i.e. 0.039 versus 0.043, which means that a large part of the initial effect on list prices remains after the transaction process.

Row (3) takes time on the market, defined as days from when an advertisement is posted online to when it is removed, as the dependent variable. Row (4) uses the total number of page views, an indicator for the attractiveness of properties to potential buyers, as the dependent variable. This variable is another proxy for the probability of sale: the more page views by potential buyers, the more likely a quick sale. The message is consistent in row (3) and row (4): the two groups show no significant difference after controlling for list price formats. The coefficient of time-on-market is small, suggesting that we can rule out substantial positive or negative total effects. Similarly, the coefficient of total page

[9] WhenFresh/Zoopla data are available from the Consumer Data Research Centre (CDRC).

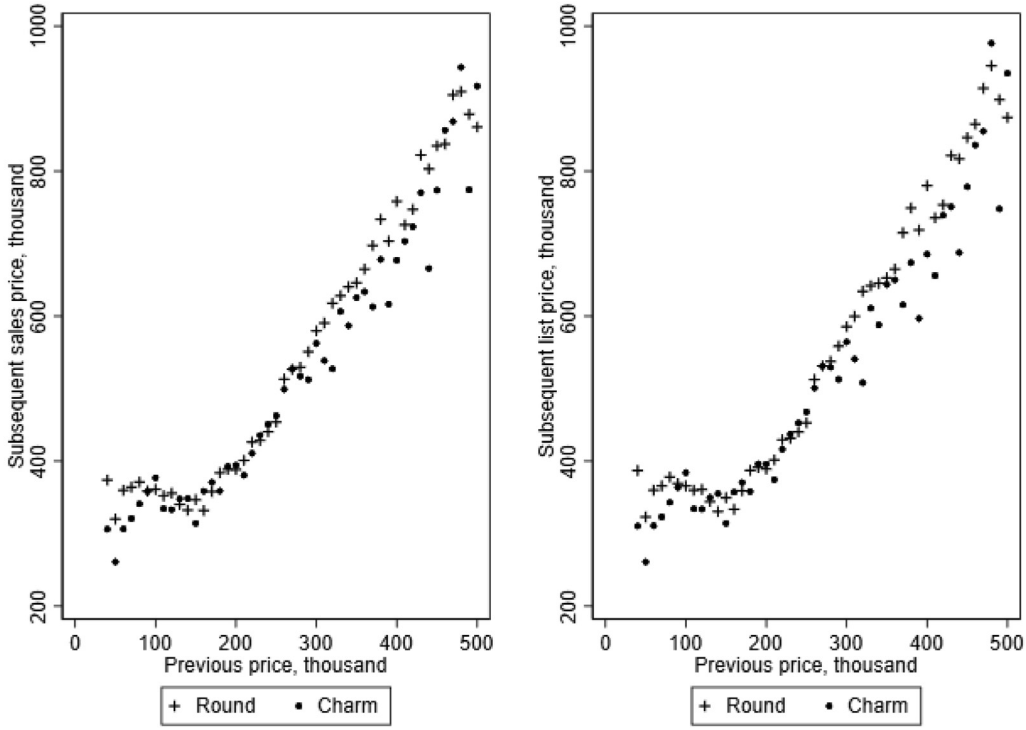


Fig. 8. Discontinuity in List Prices
Notes: This figure replicates Fig. 3(b) with the sub-sample for which list prices are available from Zoopla from 2014 to 2019. Prices are in 1,000 units. The left figure illustrates the average subsequent sales price against each previous price bin of a £1000 bin width at round number thresholds (i.e. 10,000 multiples) in pairs; the right figure illustrates the average list price.

Table 7
Effects on sales outcomes.

Dependent Variable	No. of Observations	Model 1 Coef. Round	Model 2 Coef. Round
ln(sales price - previous price)	43,220	0.039***	0.039***
		(0.003)	(0.003)
ln(list price – previous price)	43,220	0.043***	0.043***
		(0.003)	(0.003)
Time on the market	30,876	1.223	1.562
		(2.992)	(2.989)
Total page views	42,388	35.618**	34.663
		(18.105)	(18.108)
Property characteristics		√	√
Fixed effects		√	√
Charm list price			√

Notes: These are a set of regressions with different dependent variables (shown in column one). In addition to the variable of interest, i.e. Round, model 1 includes property characteristics and fixed effects; model 2 adds an additional dummy variable that equals one if the list price is a charm number.
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

views is also small relative to its mean value (35 versus 1,067). According to the second prediction in Section 2, left-digit bias affects both sellers and buyers.

8. Concluding remarks

A house transaction usually represents a household's largest single purchase, and the property also accounts for the largest proportion of their wealth. Considering the importance of the decision, I show in this research that households are inattentive to details. People make reference to the previous transaction price and are prone to left-digit bias in repeat sales transactions. A marginal previous price increase, which leads to a one-unit number change in the left-most column and a charm number upgrading to its neighbouring round number, is associated with a disproportionate increase in the reference points. The reference point increase transfers into a price premium in subsequent transactions. Therefore, the finding confirms that reference dependence and inattention have considerable influence in high-priced transactions.

The finding adds evidence for rounding behaviour, which is a cognitive heuristic that refers to the tendency of replacing a number with another value that is easy to report, remember, and communicate. Rounding is usually performed in two forms, namely, rounding-by-chop and rounding-to-the-nearest. Rounding-by-chops refers to truncating a number after certain digits. Rounding-to-the-nearest refers to rounding to the nearest number that is easy to cognitively process. For example, £234,950 will be rounded to £234,900 or £234,000 based on the first

Table A.1
Data used for LTV estimate.

Data	Geographical Unit	Time Freq.	Source
Outstanding Loans	Postcode Sector (e.g. SW9 6)	Quarterly	Council of Mortgage Lenders (CML)
House Value	Postcode Sector (e.g. SW9 6)	Quarterly	HM Land Registry
# of Households	Postcode District (e.g. SW9)	-	Office for National Statistics (ONS)
Homeownership Rate	National	-	English Housing Survey
% of Mortgage-Financed Purchases	National	-	UK Finance

Notes: The English Housing Survey is conducted by the Ministry of Housing, Communities & Local Government.

rule, and to £235,000 based on the second rule. Although the second rule is more accurate, it demands more computational costs. Rounding will inevitably introduce some round-off errors, and the errors accumulate when a sequence of calculations is based on the rounded value and may lead to considerable mistakes. The finding in this research suggests that the first rounding form is more prominent in dealing with house prices followed by significant pricing errors. Although I do not provide a direct proof, the prominence of rounding-by-chop can imply a market-wide pricing error in times or places where charm prices are commonly used.

Future research should further investigate interesting topics raised by this research, such as the relationship between inattention and household characteristics. Do mathematical skills moderate the effect? Are professional mathematicians or people who work with numbers (accountants, traders and cashiers) better at overcoming left-digit bias? Are first-time buyers more attentive than home movers?

Declaration of Competing Interest

None.

CRediT authorship contribution statement

Charlotte C. Meng: Conceptualization, Methodology, Software, Data curation, Formal analysis, Writing – original draft, Writing – review & editing.

Acknowledgements

I thank Cristian Badarinza, Michelle Baddeley, Helen Bao, Itzhak Ben-David, Phillipe Bracke, John Forth, John Glascock, Colin Lizieri, Thies Lindental, Carolin Schmidt, Vincent Yao, seminar participants at the University of Cambridge, two anonymous referees and the Editor, Edward Glaeser, for helpful comments and suggestions.

The research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

The data for this research have been provided by the Consumer Data Research Centre, an ESRC Data Investment, under project ID CDRC 373, ES/L011840/1; ES/L011891/1.

Appendix A. An estimate for the equity constraint

The core data about mortgages are the outstanding loan levels from the postcode lending scheme, part of a joint data reporting exercise co-ordinated by the British Bankers' Association (BBA) and the Council of Mortgage Lenders (CML). The scheme requires participating lenders to report the total amount of borrowing outstanding on customer accounts at the end of each quarter at UK postcode sector levels[10]. The number represents stock levels comprising borrowing agreements made in the past, new agreements, repayments and borrowing written off, rather than current demand. Reports are made on three separate lending streams: SMEs, residential mortgages and unsecured personal loans. The residential mortgage stream is used in this research. Collectively the participating lenders account for about 73 percent of mortgage lending in the UK.

The loan-to-value (LTV) estimate is the ratio between the household loan level and the home value. For the household loan level, I divide the outstanding loan level at quarter-postcode sector level by the number of households that have a mortgage in the same postcode sector. This household number is the total number of households within the postcode sector, multiply the homeownership rate and the percentage of homes bought with a mortgage. The denominator, home value, is transaction-based. It is the average house price for each quarter-postcode sector.

Hence, the LTV estimate at postcode sector p, quarter q is obtained with the following equations.

$$LTV_{p,q} = \frac{outstanding\ loan_{p,q} / \#\ of\ households\ with\ mortgages_p}{average\ home\ price_{q,p}}$$

$$\#\ of\ households\ with\ mortgages_p = \#\ of\ households_p * homeownership\ rate * \%\ of\ homes\ bought\ with\ mortgage$$

Table A.1 describes the data and their source. Outstanding loans and house values vary across time and postcode sector; the number of households varies graphically; the homeownership rate and the percentage of mortgage-financed purchases are constant for all sectors, i.e. 64 percent and 75.9 percent respectively.

[10] UK postcodes are alphanumeric and are variable in length: ranging from six to eight characters (including a space). Each postcode comprise two parts separated by a single space: the outward code and the inward code. The outward code includes an area code and a district code. The inward code includes a sector code and a unit code. Examples of full postcodes include 'SW1P 1QW', 'E20 2ST', and 'SL4 1NJ'. The postcode sector in this research includes the area code, district code and the sector code. Examples of postcode sector include 'SW1P 1', 'E20 2' and 'SL4 1'. As of June 2016, there are 11,192 postcode sectors in UK.

B. Tables and Figures

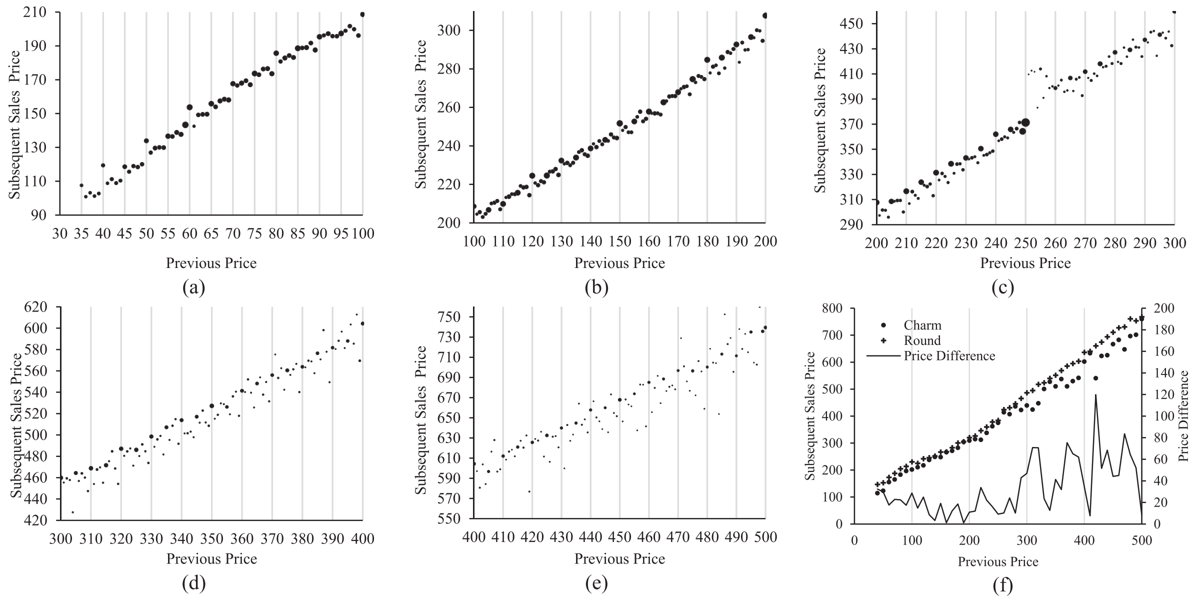


Fig. B.1. Graphical Presentations of Discontinuity (£35,000 - £500,000)
Notes: Previous prices (x-axis) and subsequent sales price (y-axis) are in 1,000 units. (a)–(e) illustrates the average subsequent price in each previous price bin of a £1000 bin width, and the dot sizes are proportional to the number of properties in the bins. (f) presents the average price (left axis) from (a)–(d) in pairs; each pair neighbours a £10,000 multiple. It also plots the discontinuity (right axis)

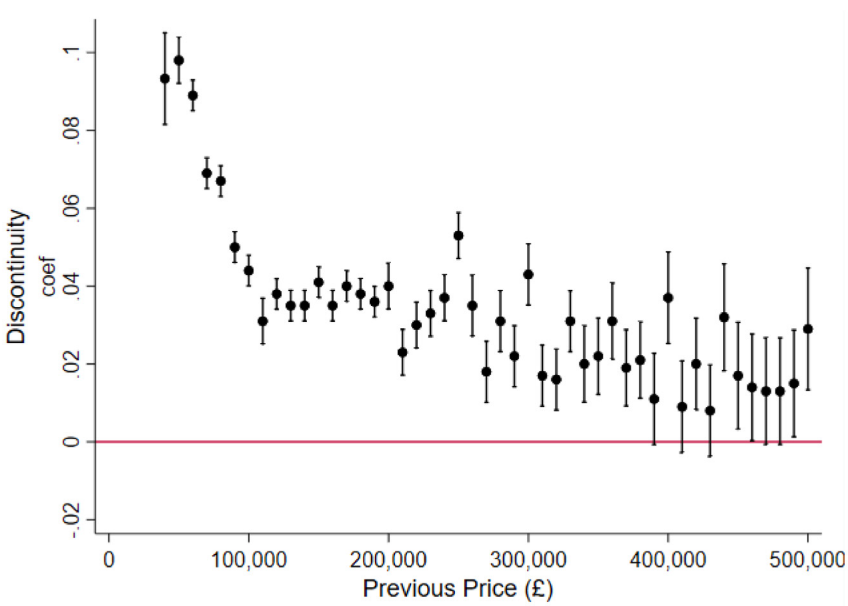


Fig. B.2. Estimates from Regression Discontinuity Design with Holding Time Control
Notes: This figure provides estimates using similar method as in Fig. 4, but with holding time (in days) as an additional control and two ways fixed effects of sales year and month. All the regressions are OLS estimates, and standard errors are clustered at the property level. The dependent variable is the log of the house price. The independent variables include a flexible function of the log of the previous price, a set of dummy variables for whether the previous price crosses a round number threshold, property characteristics, district fixed effects, year fixed effects. Standard errors are shown in parentheses. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table B.1
Balance of covariates tests.

	(1) garden	(2) patio	(3) school	(4) freehold	(5) Type = flat	(6) no of bedrooms
Round	0.002	0.001	-0.000	-0.000	0.002	0.076
	(0.003)	(0.001)	(0.001)	(0.003)	(0.003)	(0.005)
R-Squared	0.003	0.001	0.001	0.035	0.034	0.107

Notes: This table regress each property characteristic on the variable Round. Standard errors in parentheses

Table B.2
Adding holding time (in days) as an additional control.

	Full Sample (1)	(2)	Doughnut Sample (3)	(4)
Round	0.042***	0.042***	0.045***	0.042***
	(0.001)	(0.002)	(0.001)	(0.002)
Round × 5k-multiple		0.002		0.001
		(0.003)		(0.003)
Round × 10k-multiple		0.003		0.001
		(0.002)		(0.002)
Round × 50k-multiple		-0.015***		0.009
		(0.003)		(0.005)
Round × 100k-multiple		0.013**		0.026***
		(0.004)		(0.005)
Property characteristics	√	√	√	√
Fixed effects				
Magnitude	√	√	√	√
District × sale year	√	√	√	√
Sale year × month	√	√	√	√
Observations	300,007	300,007	282,251	282,251
R-Squared	0.822	0.822	0.824	0.824

Notes: This table provides the estimates based on Eq. (2) on the full sample and doughnut sample that excludes transactions around effective tax thresholds. All the regressions are OLS estimates, and standard errors are clustered at the property level. The dependent variable is the log of the difference between the current house price and the prior house price. The independent variables include the dummy variable Round which equals to one if the previous purchase price is a multiple of 1,000, property characteristics, district fixed effects, previous and current year fixed effects, magnitude of previous price fixed effects and two-way fixed effects of districts and sales year (previous and current). Standard errors are shown in parentheses. Holding time less than 6 months is excluded.
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table B.3
Allowing Heterogenous Slopes on Either Side of Round Number Thresholds To allow for heterogeneous slopes at either side of a round number, Eq. (2) is adapted as follows, $\boldsymbol{R_j}$ being the closest round number to the previous price.
$\ln P_{it} - \ln P_{is} = \ln(P_{it}/P_{is})$
$= \alpha_0 + \rho_j Round + \theta_j (P_{is} - R_j) + \gamma_j Round \times (P_{is} - R_j) + X_i'\beta + \delta_s + \delta_t + \varphi_j + e_{ist}$’

	Full Sample (1)	(2)	Doughnut Sample (3)	(4)
Round	0.058***	0.057***	0.064***	0.061***
	(0.002)	(0.002)	(0.002)	(0.002)
Round × 5k-multiple		0.003		0.003
		(0.003)		(0.003)
Round × 10k-multiple		0.003		0.002
		(0.002)		(0.003)
Round × 50k-multiple		-0.016***		0.011*
		(0.003)		(0.005)
Round × 100k-multiple		0.015***		0.028***
		(0.004)		(0.005)
Property characteristics	√	√	√	√
Fixed effects				
Magnitude	√	√	√	√
District × sale year	√	√	√	√
Sale year × month	√	√	√	√
Observations	312,138	312,138	293,722	293,722
R-Squared	0.818	0.818	0.820	0.820

Notes: This table provides the estimates based on Eq. (2) on the full sample and doughnut sample that excludes transactions around effective tax thresholds. All the regressions are OLS estimates, and standard errors are clustered at the property level. The dependent variable is the log of the difference between the current house price and the prior house price. The independent variables include the dummy variable Round which equals to one if the previous purchase price is a multiple of 1,000, property characteristics, district fixed effects, previous and current year fixed effects, magnitude of previous price fixed effects and two-way fixed effects of districts and sales year (previous and current). Standard errors are shown in parentheses. Holding time less than 6 months is excluded.
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table B.4
Residential property tax rates.

Sales price (£1000)	16 Mar 1993 - 7 July 1997	8 July 1997 - 23 Mar 1998	24 Mar 1998 - 15 Mar 1999	16 Mar 1999 - 27 Mar 2000	28 Mar 2000 - 16 Mar 2005	17 Mar 2005 - 22 Mar 2006	23 Mar 2006 - 2 Sep 2008	3 Sep 2008 - 31 Dec 2009	1 Jan 2010 - 5 Apr 2011	6 Apr 2011 - 21 Mar 2012	22 Mar 2012 - 2 Dec 2014
0-60	0	0	0	0	0	0	0	0	0	0	0
60-120	1	1	1	1	1						
120-125						1					
125-175							1		1	1	1
175-250								1			
250-500		1.5	2	2.5	3	3	3	3	3	3	3
500-1000		2	3	3.5	4	4	4	4	4	4	4
1000-2000										5	5
>2000											7
CAS											

References

Agarwal, S., Ben-David, I., Yao, V., 2015. Collateral valuation and borrower financial constraints: evidence from the residential real estate market. Manage. Sci. 61, 2220–2240.

Allen, E.J., Dechow, P.M., Pope, D.G., Wu, G., 2017. Reference-dependent preferences: evidence from marathon runners. Manage. Sci. 63, 1657–1672.

Allen, M.T., Dare, W.H., 2004. The effects of charm listing prices on house transaction prices. Real Estate Econ. 32, 695–713.

Anenberg, E., 2011. Loss aversion, equity constraints and seller behavior in the real estate market. Reg. Sci. Urban. Econ. 41, 67–76.

Arbel, Y., Ben-Shahar, D., Gabriel, S., 2014. Anchoring and housing choice: results of a natural policy experiment. Reg. Sci. Urban. Econ. 49, 68–83.

Athey, S., Imbens, G.W., 2017. The state of applied econometrics: causality and policy evaluation. J. Econ. Perspect. 31, 3–32.

Ben-David, I., 2011. Financial constraints and inflated home prices during the real estate boom. Am. Econ. J.-Appl. Econ. 3, 55–87.

Best, M.C., Kleven, H.J., 2018. Housing market responses to transaction taxes: evidence from notches and stimulus in the UK. Rev. Econ. Stud. 85, 157–193.

Bokhari, S., Geltner, D., 2011. Loss aversion and anchoring in commercial real estate pricing: empirical evidence and price index implications. Real Estate Econ. 39, 635–670.

Bracke, P., Pinchbeck, E.W., Wyatt, J., 2018. The time value of housing: historical evidence on discount rates. Econ. J. 128, 1820–1843.

Brown, J., Hossain, T., Morgan, J., 2010. Shrouded attributes and information suppression: evidence from the field. Q. J. Econ. 125, 859–876.

Bucchianeri, G.W., Minson, J.A., 2013. A homeowner's dilemma: anchoring in residential real estate transactions. J. Econ. Behav. Organ. 89, 76–92.

Busse, M.R., Lacetera, N., Pope, D.G., Silva-Risso, J., Sydnor, J.R., 2013. Estimating the effect of salience in wholesale and retail car markets. Am. Econ. Rev. 103, 575–579.

Chava, Sudheer, & Yao, Vincent W. (2017)."Cognitive reference points, the left-digit effect, and clustering in housing markets,"

Chetty, R., Looney, A., Kroft, K., 2009. Salience and taxation: theory and evidence. Am. Econ. Rev. 99, 1145–1177.

DellaVigna, S., 2009. Psychology and economics: evidence from the field. J. Econ. Lit. 47, 315–372.

DellaVigna, Stefano, Pollet, Joshua M., 2009. Investor inattention and Friday earnings announcements. J. Finance 64, 709–749.

Engelhardt, G.V., 2003. Nominal loss aversion, housing equity constraints, and household mobility: evidence from the United States. J. Urban Econ. 53, 171–195.

Erdem, T., Keane, M.P., Sun, B., 2008. A dynamic model of brand choice when price and advertising signal product quality. Market. Sci. 27, 1111–1125.

Fluet, C., Garella, P.G., 2002. Advertising and prices as signals of quality in a regime of price rivalry. Int. J. Ind Organ. 20, 907–930.

Freeman, David J., 2019. Expectations-based reference-dependence and choice under risk. Econ. J. 129, 2424–2458.

Genesove, D., Mayer, C., 2001. Loss aversion and seller behavior: evidence from the housing market. Q. J. Econ. 116, 1233–1260.

Genesove, D., Mayer, C.J., 1997. Equity and time to sale in the real estate market. Am. Econ. Rev. 87, 255–269.

Giglio, S., Maggiori, M., Stroebel, J., 2015. Very long-run discount rates. Q. J. Econ. 130, 1–53.

Hilber, C.A.L., Lyytikainen, T., 2017. Transfer taxes and household mobility: distortion on the housing or labor market? J. Urban Econ. 101, 57–73.

Holdershaw, Judith, Gendall, Philip, Garland, Ron, 2017. The widespread use of odd pricing in the retail sector. Market. Bull. 8, 53–58.

Hukkanen, Petri, Keloharju, Matti, 2019. Initial offer precision and M&A outcomes. Financ. Manag. 48, 291–310.

Jin, Ginger Zhe, Luca, Michael, Martin, Daniel, 2015. Is No News (perceived as) Bad News? An Experimental Investigation of Information Disclosure. National Bureau of Economic Research.

Kahneman, D., Tversky, A., 1979. Prospect theory - analysis of decision under risk. Econometrica 47, 263–291.

Kahneman, D., Knetsch, J.L.&, Thaler, R.H., 1990. Experimental tests of the endowment effect and the Coase theorem. J. Polit. Econ. 98, 1325–1348.

Kahneman, D.&, Tversky, A., 1979. Prospect theory - analysis of decision under risk. Econometrica 47, 263–291.

Korvorst, M., Damian, M.F., 2008. The differential influence of decades and units on multidigit number comparison. Q. J. Exp. Psychol. 61, 1250–1264.

Koszegi, B.&, Rabin, M., 2006. A model of reference-dependent preferences. Q. J. Econ. 121, 1133–1165.

Lacetera, N., Pope, D.G., Sydnor, J.R., 2012. Heuristic thinking and limited attention in the car market. Am. Econ. Rev. 102, 2206–2236.

Lee, D.S., Lemieux, T., 2010. Regression discontinuity designs in economics. J. Econ. Lit. 48, 281–355.

Levitt, S.D., Syverson, C., 2008. Market distortions when agents are better informed: the value of information in real estate transactions. Rev. Econ. Stat. 90, 599–611.

Liu, Y., Gallimore, P., Wiley, J.A., 2015. Nonlocal office investors: anchored by their markets and impaired by their distance. J. Real Estate Finance Econ. 50, 129–149.

Lynn, M., Flynn, S.M., Helion, C., 2013. Do consumers prefer round prices? Evidence from pay-what-you-want decisions and self-pumped gasoline purchases. J. Econ. Psychol. 36, 96–102.

Maltz, Amnon, 2019. Exogenous endowment—endogenous reference point. Econ. J. 130, 160–182.

Merlo, Antonio, Ortalo-Magne, Francois, 2004. Bargaining over residential real estate: evidence from England. J. Urban Econ. 56, 192–216.

Mian, A., Sufi, A., 2009. The consequences of mortgage credit expansion: evidence from the us mortgage default crisis. Q. J. Econ. 124, 1449–1496.

Poltrock, S.E., Schwartz, D.R., 1984. Comparative judgments of multidigit numbers. J. Exp. Psychol.-Learn. Memory Cogn. 10, 32–45.
Pope, D.G., Pope, J.C., Sydnor, J.R., 2015. Focal points and bargaining in housing markets. Games Econ. Behav. 93, 89–107.
Pope, D., Simonsohn, U., 2011. Round numbers as goals: evidence from baseball, sat takers, and the lab. Psychol. Sci. 22, 71–79.
Repetto, Luca, Solís, Alex, 2020. The price of inattention: evidence from the Swedish housing market. J. Eur. Econ. Assoc. 18, 3261–3304.
Rosen, S., 1974. Hedonic prices and implicit markets - product differentiation in pure competition. J. Polit. Econ. 82, 34–55.
Shlain, Avner S., 2018. More than a Penny's Worth: Left-Digit Bias and Firm Pricing. University of California, Berkeley.
Stein, J.C., 1995. Prices and trading volume in the housing-market - a model with down–payment effects. Q. J. Econ. 110, 379–406.
Stiving, M., 2000. Price-endings when prices signal quality. Manage. Sci. 46, 1617–1629.
Thaler, R., 1980. Toward a positive theory of consumer choice. J. Econ. Behav. Organ. 1, 39–60.
Thomas, M., Morwitz, V., 2005. Penny wise and pound foolish: the left-digit effect in price cognition. J. Consum. Res. 32, 54–64.
Thomas, M., Simon, D.H., Kadiyali, V., 2010. The price precision effect: evidence from laboratory and market data. Market. Sci. 29, 175–190.
Tversky, A., Kahneman, D., 1974. Judgment under uncertainty - heuristics and biases. Science 185, 1124–1131.
Tversky, A., Kahneman, D., 1992. Advances in prospect-theory - cumulative representation of uncertainty. J. Risk Uncertainty 5, 297–323.
Wiltermuth, Scott S., Gubler, Timothy, & Pierce, Lamar (2021) "Anchoring on historical round number reference points: expert experience corrects for pricing bias while market forces do not," Available at SSRN 3653476,.